รายงานการวิจัยฉบับสมบูรณ์

โครงการ “ระบบการแนะนำสถานที่ภายใต้การวิเคราะห์พฤติกรรมของผู้ใช้ (Location-based recommendation system based on user behavior analysis)”

คณะผู้วิจัย

นายโกเมศ	อัมพวัน	หัวหน้าโครงการ
นายสุเมธ	ดาราพิสุทธิ์	ผู้ช่วยวิจัย
นายณัฐนนท์	ลีลาตระกูล	ผู้ร่วมวิจัย

โครงการวิจัยประเภทงบประมาณเงินรายได้

ปีงบประมาณ พ.ศ. 2563

สัญญาเลขที่ 2/2563

รายงานวิจัยฉบับสมบูรณ์

ระบบการแนะนำสถานที่ภายใต้การวิเคราะห์พฤติกรรมของผู้ใช้

นายโกเมศ อัมพวัน
คณะวิทยาการสารสนเทศ
กันยายน 2563

บทคัดย่อ

เครือข่ายสังคมออนไลน์ทางด้านสถานที่เป็นแอปพลิเคชันหนึ่งที่ผู้ใช้มีการแบ่งปันสถานที่ร่วมกับการแสดงความคิดเห็นและข้อเสนอแนะให้กับผู้ใช้คนอื่น ๆ ภายในระบบ ณ ปัจจุบันมีหลายแอปพลิเคชันที่ได้รับความนิยมอย่างมาก เช่น Foursquare Facebook place และ Yelp เป็นต้น ในแต่ละแอปพลิเคชันจะจัดเก็บข้อมูลในระบบเป็นจำนวนมากส่งผลให้การแนะนำสถานที่ที่คาดว่าผู้ใช้ชื่นชอบเป็นเรื่องที่มีความท้าทาย โดยเฉพาะอย่างยิ่งการสร้างรายการแนะนำสถานที่ให้กับผู้ใช้ใหม่หรือผู้ใช้ที่มีข้อมูลการเช็คอินน้อยครั้ง ดังนั้น ในหลายงานวิจัยได้มีการค้นหาสถานที่ที่มีความน่าสนใจเฉพาะพื้นที่โดยพิจารณาจากการเยี่ยมชมของผู้ใช้ในพื้นที่นั้น ๆ แต่อย่างไรก็ตาม งานวิจัยก่อนหน้ามีการพิจารณาความถี่ในการเช็คอินของผู้ใช้และสถานที่เพียงมุมมองเดียวเท่านั้น อีกทั้งหลายงานวิจัยยังคงประสบปัญหาการแนะนำสถานที่ให้กับผู้ใช้ใหม่ที่ซึ่งจะส่งผลให้ความถูกต้องของการแนะนำสถานที่ให้กับผู้ใช้ลดลง ด้วยเหตุนี้ ในงานวิจัยนี้จึงได้นำเสนอวิธีการสร้างรายการแนะนำสถานที่ที่เรียกว่า “ระบบแนะนำสถานที่ที่น่าสนใจมากที่สุดเอ็นอันดับ” (N-most Interesting location-based recommender system, NILR) สำหรับแนะนำให้กับผู้ใช้ใหม่ที่ไม่มีความเชี่ยวชาญในพื้นที่ โดยวิธีการที่นำเสนอได้มีการค้นหาสถานที่ที่น่าสนใจด้วยการพิจารณาความถี่ในการเช็คอินของผู้ใช้ทั้งหมดร่วมกับการพิจารณาค่าความชอบของผู้ใช้ ด้วยคำนึงถึงการที่ผู้ใช้คนหนึ่ง ๆ กลับมาเยี่ยมชมหรือใช้บริการซ้ำ ณ สถานที่หนึ่ง ๆ ร่วมกับความถี่ในการเช็คอินของผู้ใช้ ที่ซึ่งวิธีการที่นำเสนอประกอบไปด้วย 3 ขั้นตอนหลัก คือ 1. การคำนวณค่าคะแนนความน่าสนใจของสถานที่ 2. การจัดอันดับสถานที่ที่น่าสนใจ และ 3. การสร้างรายการแนะนำสถานที่ที่น่าสนใจ จากผลการทดลองในฐานข้อมูล Foursquare แสดงให้เห็นว่าวิธีการที่นำเสนอมีความถูกต้องในการสร้างรายการแนะนำสถานที่และโดยเฉพาะความถูกต้องในการจัดอันดับที่ดีกว่าขั้นตอนวิธีการ HITS แบบดั้งเดิม

Abstract

The location-based social networks (LBSNs) are popular platform that appeal many users to share interesting locations with opinions to other users. The applications collect large sets of data including user profiles, location profiles and relationships between users and locations, making location recommendation more challenging. Especially, if users have few or no check-in histories as a new user. To cope with such scenarios, several works attempt to discover interesting locations regarding visiting of other users in interesting area. However, there exists no work that take into account both locations' visiting frequency and users' return rate to discover interesting locations. Additionally, several previous works also suffer from the cold-start problem where new users have no check-in or few check-in history causing recommenders hardly suggest locations matching users' preference. In this work, we propose an enhanced method, called N-most interesting location-based recommender system (NILR), which effectively recommends the N-most preferred places for new users divided by 3 procedures 1. the NILR discovers interesting locations by taking into account both check-in frequencies and number of return visits of previous users already in the system. 2. a ranking procedure is introduced to create the final list of interesting locations used for recommendations. Finally, 3. interesting locations are filtered again based on current location of new user. The results of the experiments on Foursquare dataset reveal that our proposed location recommender system and raking method perform effectively and efficiently, and outperform the HITS based model in terms of accuracies and rankings.

สารบัญ

หน้า

บทคัดย่อ ..... ก
Abstract ..... ข
บทที่ 1 บทนำ 1
1.1 ที่มาและความสำคัญ 1
1.2 วัตถุประสงค์ของงานวิจัย 3
1.3 ประโยชน์ที่คาดว่าจะได้รับ 4
1.4 ขอบเขตของงานวิจัย 4
บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง 5
2.1 ระบบแนะนำสถานที่ 5
2.2 การสร้างรายการแนะนำสถานที่โดยพิจารณาประวัติการเยี่ยมชมของผู้ใช้ 6
2.2 การสร้างรายการแนะนำสถานที่โดยพิจารณาผู้ใช้ภายในพื้นที่นั้น ๆ 9
2.4 ฐานข้อมูลที่ใช้ในการทดลอง 12
บทที่ 3 วิธีการที่นำเสนอ 14
3.1 นิยามพื้นฐาน 14
3.2 ระบบแนะนำสถานที่ที่น่าสนใจมากที่สุดเอ็นอันดับ 15
3.3 ตัวอย่างการสร้างรายการแนะนำสถานที่ที่น่าสนใจด้วยวิธี NILR 22
บทที่ 4 ผลการดำเนินงาน 30
4.1 การตั้งค่าการทดลองและวิธีการที่ใช้ในการวัดประสิทธิภาพความถูกต้อง 30
4.2 ผลการทดลอง 32
บทที่ 5 สรุปและอภิปรายผล 40
บรรณานุกรม 41

สารบัญตาราง

หน้า

ตารางที่ 2.1 ปัจจัยที่ใช้ในการพิจารณาเพื่อค้นหาสถานที่ที่น่าสนใจ 12
ตารางที่ 2.2 คุณลักษณะของเมืองต่าง ๆ ที่ถูกใช้ในการทดลอง..... 13

สารบัญภาพ

หน้า

ภาพที่ 2.1 ตัวอย่างของเครือข่ายสังคมออนไลน์ทางด้านสถานที่ 5
ภาพที่ 2.2 ขั้นตอนวิธีการ Hyperlink-Induced Topic Search (HITS)..... 9
ภาพที่ 2.3 ตัวอย่างของฐานข้อมูล Foursquare ที่ใช้ในการทดลอง..... 12
ภาพที่ 3.1 ขั้นตอนวิธีการสร้างรายการแนะนำสถานที่ NILR..... 15
ภาพที่ 3.2 ขั้นตอนวิธี HITS สำหรับคำนวณค่าคะแนนความน่าสนใจของสถานที่โดยพิจารณาความถี่ในการเช็คอินของผู้ใช้ 17
ภาพที่ 3.3 ขั้นตอนวิธี ReturnHITS สำหรับคำนวณค่าคะแนนความน่าสนใจของสถานที่โดยพิจารณาค่าความชอบของผู้ใช้..... 18
ภาพที่ 3.4 รายละเอียดขั้นตอนวิธีการจัดอันดับรายการแนะนำสถานที่..... 20
ภาพที่ 3.5 แสดงรายละเอียดขั้นตอนวิธีการสร้างรายการแนะนำสถานที่ที่น่าสนใจมากที่สุด N อันดับ..... 21
ภาพที่ 3.6 ภาพรวมรายละเอียดขั้นตอนวิธีการสร้างรายการแนะนำสถานที่ NILR 21
ภาพที่ 3.7 ตัวอย่างข้อมูลประวัติการเช็คอินของผู้ใช้ 22
ภาพที่ 3.8 เมทริกซ์สำหรับการคำนวณค่าคะแนนความน่าสนใจของสถานที่ 22
ภาพที่ 3.9 ตัวอย่างการคำนวณค่าคะแนนความน่าสนใจของสถานที่ l_1 และค่าคะแนนผู้ใช้ u_1 โดยพิจารณาความถี่ในการเช็คอินของผู้ใช้..... 23
ภาพที่ 3.10 ตัวอย่างการปรับลดค่าคะแนน $isf(l_1)$ และ $isf(u_1)$ โดยใช้ *L2*-norm..... 24
ภาพที่ 3.11 ตัวอย่างการคำนวณค่าคะแนนความน่าสนใจของสถานที่ l_1 และค่าคะแนนของผู้ใช้เมื่ออัพเดทค่าคะแนนทั้งหมด 3 รอบ..... 25
ภาพที่ 3.12 ตัวอย่างการคำนวณค่าความชอบผู้ใช้ u_1 และค่าคะแนนความน่าสนใจสถานที่ l_1 26
ภาพที่ 3.13 ตัวอย่างการคำนวณค่าความชอบของสถานที่ l_1 และค่าคะแนนของผู้ใช้ u_1 26
ภาพที่ 3.14 ตัวอย่างการปรับลดค่าคะแนน $isp(l_1)$ และ $isp(u_1)$โดยใช้ *L2*-norm 27
ภาพที่ 3.15 ตัวอย่างค่าคะแนนของสถานที่ l_1 และค่าคะแนนของผู้ใช้โดยพิจารณาค่าความชอบเมื่อคำนวณจนครบ 3 รอบ..... 28
ภาพที่ 3.16 ตัวอย่างวิธีการจัดอันดับรายการแนะนำสถานที่ที่น่าสนใจ..... 29

สารบัญภาพ (ต่อ)

หน้า

ภาพที่ 3.17 ตัวอย่างวิธีการสร้างรายการแนะนำสถานที่ที่น่าสนใจให้กับผู้ใช้ nu 3 อันดับที่น่าสนใจที่สุด 29
ภาพที่ 4.1 ตัวอย่างการจำลองสถานที่ปัจจุบันของผู้ใช้โดยใช้วิธี MBR 31
ภาพที่ 4.2 ผลความถูกต้องโดยใช้วิธีการวัดประสิทธิภาพ Precision โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Chiyoda 32
ภาพที่ 4.3 ผลความถูกต้องโดยใช้วิธีการวัดประสิทธิภาพ Precision โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Shibuya 33
ภาพที่ 4.4 ผลความถูกต้องโดยใช้วิธีการวัดประสิทธิภาพ Precision โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Minato 33
ภาพที่ 4.5 ผลความถูกต้องโดยใช้วิธีการวัดประสิทธิภาพ Precision โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Chuo 34
ภาพที่ 4.6 ผลความถูกต้องโดยใช้วิธีการวัดประสิทธิภาพ Precision โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Shinjuku 34
ภาพที่ 4.7 ผลความถูกต้องโดยใช้วิธีการวัดประสิทธิภาพ Recall โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Chiyoda 35
ภาพที่ 4.8 ผลความถูกต้องโดยใช้วิธีการวัดประสิทธิภาพ Recall โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Shibuya 35
ภาพที่ 4.9 ผลความถูกต้องโดยใช้วิธีการวัดประสิทธิภาพ Recall โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Minato 36
ภาพที่ 4.10 ผลความถูกต้องโดยใช้วิธีการวัดประสิทธิภาพ Recall โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Shinjuku 36
ภาพที่ 4.11 ผลความถูกต้องโดยใช้วิธีการวัดประสิทธิภาพ Recall โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Chuo 37
ภาพที่ 4.12 ผลความถูกต้องของการจัดอันดับโดยใช้วิธีการวัดประสิทธิภาพ Average ranking โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Chiyoda 37
ภาพที่ 4.13 ผลความถูกต้องของการจัดอันดับโดยใช้วิธีการวัดประสิทธิภาพ Average ranking โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Shibuya 38

สารบัญภาพ (ต่อ)

หน้า

ภาพที่ 4.14 ผลความถูกต้องของการจัดอันดับโดยใช้วิธีการวัดประสิทธิภาพ Average ranking โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Minato 38
ภาพที่ 4.15 ผลความถูกต้องของการจัดอันดับโดยใช้วิธีการวัดประสิทธิภาพ Average ranking โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Shinjuku..... 39
ภาพที่ 4.16 ผลความถูกต้องของการจัดอันดับโดยใช้วิธีการวัดประสิทธิภาพ Average ranking โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Chuo 39

บทที่ 1
บทนำ

1.1 ที่มาและความสำคัญ

เครือข่ายสังคมออนไลน์ทางด้านสถานที่ (Location based social network, LBSN) ได้รับความนิยมและความแพร่หลายอย่างมากในปัจจุบัน ซึ่งมีจำนวนผู้ใช้งานที่เพิ่มสูงขึ้นอย่างรวดเร็วจากเทคโนโลยีอินเทอร์เน็ตและสมาร์ทโฟนที่ล้ำสมัย ส่งผลให้การเข้าถึงเครือข่ายสังคมออนไลน์ได้ง่ายและสะดวกมากยิ่งขึ้น (Zheng & Xie, 2011) (Zheng et al., 2014) (Yu & Chen, 2015)(Geng et al., 2019)(Xu et al., 2020) ผู้ใช้ในเครือข่ายสังคมออนไลน์ทางด้านสถานที่มักจะมีการแบ่งปันสถานที่ร่วมกับการแสดงความคิดเห็นและข้อเสนอแนะที่ผู้ใช้คนนั้นได้ไปเยี่ยมชมจริง ซึ่งมีหลายแอปพลิเคชันในเครือข่ายสังคมออนไลน์ทางด้านสถานที่ที่เป็นที่นิยม เช่น Foursquare[1] Facebook place[2] และ Yelp[3] เป็นต้น ได้ให้ผู้ใช้เช็คอิน (Check-in) และเพิ่มสถานที่ใหม่เพื่อแบ่งปันความรู้เกี่ยวกับสถานที่ที่ผู้ใช้ไปเยี่ยมชมผ่านเครือข่ายสังคมออนไลน์นี้ โดยแอปพลิเคชันดังกล่าวได้จัดเก็บข้อมูลผู้ใช้ ข้อมูลสถานที่ และความสัมพันธ์ระหว่างผู้ใช้และสถานที่ ส่งผลให้การแนะนำสถานที่จากจำนวนสถานที่ที่มีจำนวนมากที่คาดว่าผู้ใช้ชื่นชอบเป็นเรื่องที่มีความท้าทายและมีความยาก

ณ ปัจจุบันระบบแนะนำสถานที่ (Location based recommender system, LBRS) (Zhou et al., 2012)(Levandoski et al., 2012)(Ding et al., 2018) เข้ามามีบทบาทสำคัญในการช่วยการตัดสินใจ คัดกรองและแนะนำสถานที่ที่ผู้ใช้ชื่นชอบหรือคาดว่าจะเยี่ยมชม โดยประโยชน์ของระบบแนะนำสถานที่สามารถช่วยเพิ่มความพึงพอใจให้กับผู้ใช้ที่มีต่อแอปพลิเคชัน และในส่วนของเจ้าของธุรกิจสามารถใช้ประโยชน์จากระบบแนะนำเพื่อให้ผู้ใช้เข้าถึงสถานที่นั้นได้ง่ายมากขึ้น ซึ่งสามารถช่วยเพิ่มผลกำไรและช่วยเพิ่มฐานลูกค้าให้กับธุรกิจนั้น ๆ ในระบบแนะนำสถานที่ส่วนใหญ่มีการพิจารณาความชอบของผู้ใช้ที่ต้องการแนะนำจากข้อมูลของการเช็คอินหรือประวัติการเช็คอินจำนวนมาก แต่ในโลกแห่งความเป็นจริง ผู้ใช้มักจะมีการเช็คอินจำนวนน้อยครั้งและมักเยี่ยมชมสถานที่ในระแวกใกล้เคียงกับสถานที่ปัจจุบันของผู้ใช้ เช่น บ้าน อะพาร์ตเมนต์ และที่ทำงาน เป็นต้น อีกทั้งผู้ใช้ไม่สามารถเยี่ยมชมสถานที่ได้ทั้งหมด ด้วยเหตุนี้ การแนะนำสถานที่ให้กับผู้ใช้ใหม่หรือนักท่องเที่ยวจึงเป็นเรื่องที่ยาก

---

[1] https://www.foursquare.com

[2] https://www.facebook.com

[3] https://www.yelp.com

จากปัญหาดังกล่าวข้างต้น ส่งผลต่อประสิทธิภาพทางด้านความถูกต้องในการสร้างรายการแนะนำสถานที่ หลายงานวิจัยจึงได้มีการค้นหาผู้เชี่ยวชาญเฉพาะพื้นที่หรือใช้ประโยชน์จากผู้ใช้ภายในพื้นที่นั้น ๆ ซึ่งสามารถแนะนำสถานที่ที่น่าสนใจเฉพาะพื้นที่ได้และสามารถเพิ่มประสิทธิภาพความถูกต้องในการสร้างรายการแนะนำให้กับผู้ใช้ใหม่ได้ แต่อย่างไรก็ตาม ในหลายงานวิจัยก่อนหน้ามีการค้นหาสถานที่ที่น่าสนใจโดยพิจารณาเพียงความถี่ในการเช็คอินหรือเยี่ยมชมเท่านั้น (Zheng et al., 2009)(Bao et al., 2012)(Bagci & Karagoz, 2016) ซึ่งบางสถานที่ที่ถูกเยี่ยมชมจำนวนมาก แต่ผู้ใช้จำนวนมากมีการมาเยี่ยมชมเพียงครั้งเดียว (ไม่มีการกลับมาเยี่ยมชม) แสดงถึงสถานที่นั้นอาจจะเป็นสถานที่ใหม่หรือเป็นสถานที่ที่ได้รับความนิยมเพียงชั่วขณะ สำหรับสถานที่ที่มีจำนวนการเยี่ยมชมในระดับปานกลางแต่ผู้ใช้มักมีการกลับมาเยี่ยมชมเสมอนั้นแสดงถึงความนิยมแบบสม่ำเสมอ และในอีกสถานการณ์หนึ่ง ถ้าความถี่ในการเยี่ยมชมของสถานที่มีจำนวนที่เท่ากันนั้น สถานที่ที่มีการกลับมาเยี่ยมชมจากหลาย ๆ ผู้ใช้มากที่สุดแสดงถึงเป็นสถานที่ที่มีคุณภาพและได้รับความนิยมที่มากกว่าสถานที่อื่น ๆ จากการศึกษางานวิจัยก่อนหน้ายังไม่มีงานวิจัยใดที่ค้นหาสถานที่ที่น่าสนใจ โดยพิจารณาความถี่ในการเช็คอินร่วมกับการกลับมาเยี่ยมชมของผู้ใช้ และอีกหนึ่งปัญหาคือการแนะนำสถานที่ให้กับผู้ใช้ใหม่ที่เรียกว่าปัญหาการเริ่มต้นได้ยากของผู้ใช้ (user cold-start problem) เนื่องจากผู้ใช้ใหม่มักไม่มีการเช็คอินหรือมีข้อมูลการเช็คอินที่น้อยครั้ง ทำให้การค้นหาความชอบของผู้ใช้คนดังกล่าวเป็นเรื่องที่ยาก

ด้วยเหตุนี้ ในงานวิจัยนี้จึงได้นำเสนอวิธีการสร้างรายการแนะนำสถานที่ที่เรียกว่า “ระบบแนะนำสถานที่ที่น่าสนใจมากที่สุดเอ็นอันดับ” (N-most Interesting location-based recommender system, NILR) สำหรับแนะนำให้กับผู้ใช้ใหม่ที่ไม่มีความเชี่ยวชาญในพื้นที่ โดยวิธีการที่นำเสนอได้มีการค้นหาสถานที่ที่น่าสนใจด้วยการพิจารณาความถี่ในการเช็คอินของผู้ใช้ทั้งหมดร่วมกับการพิจารณาค่าความชอบของผู้ใช้ ด้วยคำนึงถึงการที่ผู้ใช้คนหนึ่ง ๆ กลับมาเยี่ยมชมหรือใช้บริการซ้ำ ณ สถานที่หนึ่ง ๆ ร่วมกับความถี่ในการเช็คอินของผู้ใช้ ที่ซึ่งวิธีการที่นำเสนอประกอบไปด้วย 3 ขั้นตอนหลัก คือ 1. การคำนวณค่าคะแนนความน่าสนใจของสถานที่ โดยการปรับปรุงขั้นตอนวิธี HITS (Hyperlink-Induced Topic Search model) (Zheng et al., 2009)(Bao et al., 2012)(Bagci & Karagoz, 2016)(Chen et al., 2016)(Ying et al., 2017) ด้วยการพิจารณาความถี่ในการเช็คอินของผู้ใช้ร่วมกับค่าความชอบของผู้ใช้จากการคำนึงถึงการกลับมาเยี่ยมชมหรือใช้บริการซ้ำต่อสถานที่ของผู้ใช้ 2. การจัดอันดับสถานที่ที่น่าสนใจ จากค่าคะแนนที่มากที่สุดของเซทรายการสถานที่ด้วยการพิจารณาความถี่ในการเช็คอินของผู้ใช้และเซทรายการสถานที่ด้วยการพิจารณาค่าความชอบของผู้ใช้โดยใช้การคำนวณแบบวนซ้ำ และ 3. การสร้างรายการแนะนำสถานที่ที่น่าสนใจ โดยระบบจะทำการเลือกสถานที่ที่น่าสนใจมากที่สุด N อันดับที่ตั้งอยู่ระแวกใกล้เคียงกับที่อยู่ปัจจุบันของผู้ใช้จากรายการแนะนำสถานที่ที่ถูกจัดอันดับไว้จากขั้นตอนก่อนหน้า ในส่วนของการ

ประเมินผลวิธีการที่นำเสนอนี้ ได้ทำการทดลองโดยใช้ฐานข้อมูลการเช็คอินของผู้ใช้จาก Foursquare (Yang et al., 2015) ใน 5 เมือง คือ Chiyoda Minato Shinjuku Shibuya และ Chuo สำหรับการจำลองที่อยู่ปัจจุบันของผู้ใช้เพื่อสร้างรายการแนะนำสถานที่ที่น่าสนใจเฉพาะพื้นที่ได้ใช้การกำหนดของเขตของผู้ใช้ที่เรียกว่า Minimum Bounding Rectangles (MBRs) (Bao et al., 2012)(Ying et al., 2017) และเปรียบเทียบกับขั้นตอนวิธี HITS แบบดั้งเดิมใน 3 วิธีการวัดประสิทธิภาพความถูกต้องคือ 1. Precision 2. Recall และ 3. Average ranking โดยวิธีการที่นำเสนอสามารถให้ความถูกต้องที่ดีกว่าขั้นตอนวิธี HITS โดยใช้วิธี Precision และ Recall ตามลำดับและมีความถูกต้องในการจัดอันดับที่ดีกว่าขั้นตอนวิธี HITS โดยใช้วิธี Average ranking ตามลำดับ

1.2 วัตถุประสงค์ของงานวิจัย

1. เพื่อพัฒนาขั้นตอนวิธีการสร้างรายการแนะนำสถานที่ที่น่าสนใจให้มีประสิทธิภาพที่เพิ่มขึ้นในด้านความถูกต้อง โดยพิจารณา 2 ปัจจัยในการสร้างรายการแนะนำสถานที่ คือ ความถี่ในการเช็คอินของผู้ใช้ร่วมกับความชอบของผู้ใช้โดยคำนึงถึงความถี่ในการเช็คอินร่วมกับการกลับมาเยี่ยมชมของผู้ใช้

2. เพื่อค้นหาสถานที่ที่น่าสนใจเฉพาะพื้นที่ อาทิเช่น ร้านอาหาร ร้านกาแฟ ร้านพิซซ่า ร้านขนมหวาน คาเฟ่ และร้านเบเกอรี่ เป็นต้น โดยการพิจารณาทั้งความถี่ในการเช็คอินร่วมกับความชอบของผู้ใช้

3. เพื่อแก้ไขปัญหาการแนะนำสถานที่ให้กับผู้ใช้ใหม่ที่พึ่งเข้าสู่ระบบและผู้ใช้ที่มีข้อมูลการเช็คอินน้อยครั้งที่ซึ่งมีข้อมูลในการเช็คอินไม่เพียงพอต่อการสร้างรายการแนะนำสถานที่แบบดั้งเดิม

4. เพื่อสร้างองค์ความรู้ใหม่ของรายการแนะนำสถานที่โดยการพิจารณาผู้ใช้ที่มักมีพฤติกรรมในการกลับมาเยี่ยมชมหรือใช้บริการซ้ำ ณ สถานที่หนึ่ง ๆ ที่ซึ่งแตกต่างจากงานวิจัยก่อนหน้า

5. เพื่อเป็นขั้นตอนวิธีพื้นฐานในการพัฒนาระบบเทคโนโลยีสารสนเทศสำหรับการแนะนำสถานที่ที่น่าสนใจให้กับผู้ใช้เพื่อช่วยในการตัดสินใจโดยการพิจารณาความถี่ในการเช็คอินของผู้ใช้ร่วมกับความชอบของผู้ใช้

6. เพื่อให้ผู้ที่มีความสนใจในแนวคิดที่นำเสนอสามารถศึกษาและประยุกต์ใช้ในงานวิจัยและประยุกต์ใช้กับระบบแนะนำสถานที่ของข้อมูลจริงได้

1.3 ประโยชน์ที่คาดว่าจะได้รับ

1. ได้ขั้นตอนวิธีการสร้างรายการแนะนำสถานที่ที่น่าสนใจที่มีความถูกต้องเพิ่มมากขึ้น โดยพิจารณา 2 ปัจจัยในการสร้างรายการแนะนำสถานที่ คือ ความถี่ในการเช็คอินของผู้ใช้โดยใช้วิธีการร่วมกับความชอบของผู้ใช้โดยคำนึงถึงความถี่ในการเช็คอินร่วมกับการกลับมาเยี่ยมชมของผู้ใช้

2. ได้องค์ความรู้ใหม่ของของรายการแนะนำสถานที่โดยการพิจารณาผู้ใช้ที่มักมีพฤติกรรมในการกลับมาเยี่ยมชมหรือใช้บริการซ้ำ ณ สถานที่หนึ่ง ๆ

3. สามารถแก้ไขปัญหาการแนะนำสถานที่ให้กับผู้ใช้ใหม่ที่พึ่งเข้าสู่ระบบและผู้ใช้ที่มีข้อมูลการเช็คอินน้อยครั้งที่ซึ่งไม่เพียงพอต่อการสร้างรายการแนะนำแบบดั้งเดิม

4. ได้ขั้นตอนวิธีพื้นฐานในการพัฒนาระบบเทคโนโลยีสารสนเทศสำหรับการแนะนำสถานที่ที่น่าสนใจให้กับผู้ใช้เพื่อช่วยในการตัดสินใจโดยการพิจารณาความถี่ในการเช็คอินของผู้ใช้ร่วมกับความชอบของผู้ใช้

5. ได้ผลงานวิจัยตีพิมพ์และเผยแพร่ในวารสารวิชาการ

1.4 ขอบเขตของงานวิจัย

1. ฐานข้อมูลที่ใช้ในการทดลองของขั้นตอนวิธีการที่นำเสนอนี้ได้ใช้การจัดเก็บประวัติการเช็คอินของผู้ใช้จากเว็บไซต์ Foursquare (Yang et al., 2015) ประกอบด้วย 1. รหัสผู้ใช้ 2. รหัสสถานที่ 3. รหัสหมวดหมู่สถานที่ 4. หมวดหมู่สถานที่ 5. ละติจูด 6. ลองจิจูด 7. เวลาในการเช็คอิน 8. เวลา UTC โดยจัดเก็บในช่วง 12 เมษายน 2012 ถึง 16 กุมภาพันธ์ 2013 (ประมาณ 10 เดือน)

2. การแนะนำสถานที่จะสามารถดำเนินการได้กับข้อมูลที่ประกอบด้วย ละติจูด (Latitude) ลองติจูด (Longitude) ของสถานที่และหมวดหมู่ของสถานที่

3. ในการทดลองของขั้นตอนวิธีการที่นำเสนอนี้ได้เลือกใช้เฉพาะหมวดหมู่อาหารซึ่งเป็นหมวดหมู่ที่ถูกเช็คอินมากที่สุดและทำการทดลองใน 5 พื้นที่ ได้แก่ Chiyoda Minato Shinjuku Shibuya และ Chuo โดยเรียงลำดับจากข้อมูลการเช็คอินมากที่สุดไปยังน้อยที่สุดตามลำดับ โดยการแปลงละติจูดและลองจิจูดไปยังพื้นที่ต่าง ๆ ด้วย Geocoding API[4]

4. การจำลองที่อยู่ปัจจุบันของผู้ใช้สำหรับกำหนดขอบเขตของการเลือกสถานที่เพื่อนำมาสร้างรายการแนะนำได้ใช้วิธีการ Minimum Bounding Rectangles (MBRs) (Ying et al., 2017)

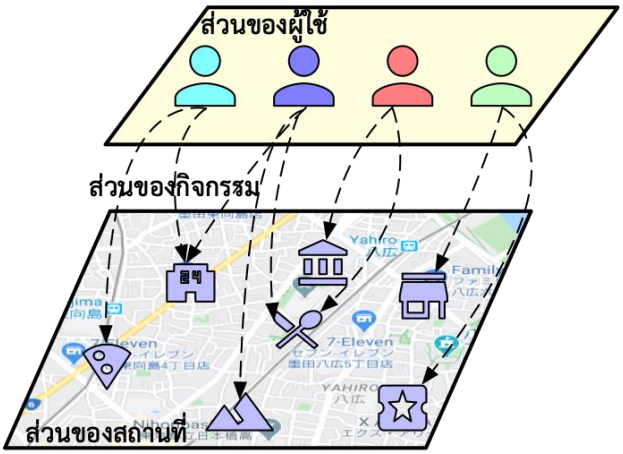
5. การประเมินผลประสิทธิภาพของวิธีการที่นำเสนอนี้ได้ใช้การวัดความถูกต้องด้วยวิธี Precision และวิธี Recall และการวัดความถูกต้องของอันดับสถานที่ในรายการแนะนำด้วยวิธี Average ranking

---

[4] https://developers.google.com/maps/documentation/geocoding/

บทที่ 2
ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงประวัติความเป็นมาและความสำคัญของระบบแนะนำสถานที่และงานวิจัยก่อนหน้าเกี่ยวกับระบบแนะนำสถานที่ งานวิจัยนี้ได้ทำการแบ่งวิธีการสร้างรายการแนะนำสถานที่ออกเป็น 2 หมวดหมู่คือ 1. การสร้างรายการแนะนำสถานที่โดยพิจารณาประวัติการเยี่ยมชมของผู้ใช้ 2. การสร้างรายการแนะนำสถานที่โดยพิจารณาจากผู้เชี่ยวชาญเฉพาะพื้นที่หรือผู้ใช้ภายในพื้นที่นั้น ๆ และสุดท้ายรายละเอียดของฐานข้อมูลที่ใช้ในการทดลองจะถูกนำเสนอตามลำดับ



ภาพที่ 2.1 ตัวอย่างของเครือข่ายสังคมออนไลน์ทางด้านสถานที่

2.1 ระบบแนะนำสถานที่

ระบบแนะนำสถานที่ (Levandoski et al., 2012)(Yu, Y. & Chen, X., 2015)(Ding et al., 2018)(Xu et al., 2020) ได้รับความนิยมอย่างมากในปัจจุบัน เนื่องจากเทคโนโลยีอินเทอร์เน็ตและสมาร์ทโฟนที่ล้ำสมัย ผู้ใช้มักมีการแบ่งปันสถานที่ที่ชื่นชอบร่วมกับการแสดงความคิดเห็นและข้อเสนอแนะกับผู้ใช้คนอื่น ๆ ในระบบ เช่น ร้านอาหาร ห้างสรรพสินค้า ร้านกาแฟ ร้านเบเกอรี่ และร้านสะดวกซื้อ เป็นต้น จากรายงานในเดือนมิถุนายน ปีค.ศ. 2016 แอปพลิเคชัน Foursqaure ได้มีการจัดเก็บสถิติในการเช็คอินของผู้ใช้ซึ่งมีมากกว่า 8 พันล้านเช็คอิน และสถานที่ที่ผู้ใช้ทำการเช็คอินมากกว่า 65 ล้านสถานที่รวมถึงผู้ใช้ในระบบซึ่งมากกว่า 55 ล้านผู้ใช้ในแต่ละเดือน (Zhao & Lyu, 2016) และจากรายงานของแอปพลิเคชัน Yelp เปิดเผยว่ามีผู้ใช้งานต่อเดือนประมาณ 29 ล้านผู้ใช้

(Xu et al., 2020) จากข้อมูลที่กล่าวมาแสดงให้เห็นถึงความนิยมและความสนใจจากผู้ใช้จำนวนมาก ส่งผลให้ระบบแนะนำสถานที่ (Location based recommender system) ได้เข้ามามีบทบาทสำคัญในการช่วยคัดกรองและแนะนำสถานที่ที่ผู้ใช้ชื่นชอบหรือคาดว่าจะเยี่ยมชม ซึ่งระบบแนะนำสถานที่สามารถแบ่งการแนะนำออกเป็น 3 ส่วน คือ 1. การแนะนำผู้ใช้ 2. การแนะนำกิจกรรม 3. การแนะนำสถานที่ ดังภาพที่ 2.1 สำหรับการแนะนำผู้ใช้นั้น ผู้ใช้ในระบบจะถูกพิจารณาความชอบที่คล้ายกับผู้ใช้ที่ร้องขอรายการแนะนำ เช่น กลุ่มของเพื่อนหรือผู้เชี่ยวชาญในพื้นที่ เป็นต้น เพื่อสร้างรายการแนะนำสถานที่และกลุ่มของผู้ใช้ดังกล่าวสามารถที่จะสร้างเป็นรายการแนะนำผู้ใช้ได้ (Xiao et al., 2010) (Symeonidis et al., 2011)(Wang et al., 2014) สำหรับการแนะนำกิจกรรมจะเป็นการแนะนำกิจกรรมที่ขึ้นกับช่วงเวลา เช่น การเลือกซื้อของที่ห้างสรรพสินค้าในช่วงบ่ายหรือวันหยุดและการเลือกดูหนังในช่วงเวลากลางคืน เป็นต้น (Zheng et al., 2010a) (Zheng et al., 2010b) (Yang et al, 2015) สำหรับการแนะนำสถานที่ ประกอบด้วย การแนะนำสถานที่เดียวให้กับผู้ใช้ และการแนะนำลำดับของสถานที่ให้กับผู้ใช้หรือทริป (Yuan et al., 2013) (Lim et al., 2018)(Lim et al., 2019)

2.2 การสร้างรายการแนะนำสถานที่โดยพิจารณาประวัติการเยี่ยมชมของผู้ใช้

ในปี 2013 Yuan และคณะ (Yuan, 2013) จึงได้นำเสนอขั้นตอนวิธีการสร้างรายการแนะนำสถานที่ให้กับผู้ใช้ โดยการประยุกต์ใช้วิธีการกรองร่วมร่วมกับพิจารณาผู้ใช้เป็นหลัก นอกจากนี้วิธีการนี้ยังคำนึงถึงช่วงเวลาในการเช็คอินของผู้ใช้ที่เรียกว่า time-aware POI recommendation อีกทั้งยังแก้ไขปัญหาการลดความเบาบางของข้อมูล (Sparsity problem) ด้วยเทคนิค Smooting ซึ่งคือการพิจารณาชั่วโมงข้างเคียงในการเช็คอินของผู้ใช้และพิจารณาการเช็คอินของสถานที่จากผู้ใช้ข้างเคียงที่มีความคล้ายกับผู้ใช้เป้าหมายและใช้เทคนิค Bayes rules ในการวิเคราะห์พฤติกรรมในการเช็คอินของผู้ใช้เพื่อสร้างรายการแนะนำสถานที่ จากขั้นตอนวิธีการที่นำเสนอนี้สามารถเพิ่มความถูกต้องในการแนะนำสถานที่ให้กับผู้ใช้และลดปัญหาความเบาบางของข้อมูลได้

Liu และคณะ (Liu et al., 2013) ได้นำเสนอขั้นตอนวิธีการสร้างรายการแนะนำสถานที่ที่คาดว่าผู้ใช้จะทำการเยี่ยมชมต่อไปโดยการใช้เทคนิคการแยกเมทริกซ์ (Matrix factorization) ในการทำนายค่าคะแนนความชอบของผู้ใช้จากการพิจารณารูปแบบความชอบของผู้ใช้จากหมวดหมู่ (Category) ที่ผู้ใช้ทำการเช็คอินและทำการจัดกลุ่มผู้ใช้ที่ทำการเช็คอินตามช่วงเวลาร่วมกับการคำนึงถึงสถานที่ปัจจุบันของผู้ใช้โดยใช้คุณสมบัติ power-law ในการสร้างรายการแนะนำ จากขั้นตอนของวิธีการที่นำเสนอนี้สามารถสร้างรายการแนะนำได้อย่างถูกต้องและแก้ไขปัญหาความเบาบางของข้อมูลโดยใช้การจัดกลุ่มของผู้ใช้อย่างมีประสิทธิภาพ

Gao และคณะ (Gao et al., 2013) นำเสนอขั้นตอนการสร้างรายการแนะนำสถานที่โดยศึกษารูปแบบพฤติกรรมการเช็คอินของผู้ใช้ในแต่ละวัน ซึ่งสามารถแบ่งได้เป็น 2 คุณลักษณะคือ 1.

Non-uniformness คือผู้ใช้ที่มีรูปแบบการเช็คอินของสถานที่ ๆ แตกต่างกันในแต่ละชั่วโมง 2. Consecutiveness คือผู้ใช้ที่มีรูปแบบการเช็คอินของสถานที่ ๆ ใกล้เคียงกันในช่วงเวลาที่ต่อเนื่องกัน ดังนั้นวิธีการนี้จึงทำการพิจารณารูปแบบสถานที่ ๆ ถูกเช็คอินในแต่ละชั่วโมงโดยใช้เทคนิคการหาความคล้ายคลึงแบบโคไซน์ (Cosine similarity) ร่วมกับใช้เทคนิคการทำนายค่าคะแนนความชอบของผู้ใช้ที่เรียกว่า low rank matrix factorization จากเทคนิคนี้สามารถสร้างรายการแนะนำได้อย่างถูกต้องและมีความเหมาะสมตามช่วงเวลา

Preotiuc-Pietro และ Cohn (Preotiuc-Pietro & Cohn, 2013) ได้มีการศึกษาพฤติกรรมและรูปแบบการเช็คอินของผู้ใช้ในฐานข้อมูล Foursquare ในหลาย ๆ มุมมอง เช่น รูปแบบการเช็คอินของผู้ใช้ซึ่งมีการเช็คอินบ่อยครั้ง ช่วงเวลาในการกลับมาสถานที่เดิม การกระจายตัวของสถานที่ ๆ ถูกเช็คอิน หมวดหมู่ของการเช็คอินของผู้ใช้ตามช่วงเวลา การจัดกลุ่มรูปแบบการเช็คอินของผู้ใช้ด้วยเทคนิค K-means เพื่อพิจารณากลุ่มของผู้ใช้ เป็นต้น อีกทั้งยังมีการศึกษาพฤติกรรมและทำนายสถานที่ต่อไปที่ผู้ใช้ต้องการเยี่ยมชมจากประวัติการณ์การเยี่ยมชมในอดีตของผู้ใช้ด้วยเทคนิค Order-K Markov อีกด้วย

Yin และคณะ (Yin et al., 2013) ได้นำเสนอขั้นตอนวิธีการสร้างรายการแนะนำสถานที่และรายการแนะนำกิจกรรมที่ควรจะทำในพื้นที่เหล่านั้น โดยนำเสนอขั้นตอนวิธีการที่เรียกว่า Location-content-aware recommender system (LCARS) ซึ่งแบ่งขั้นตอนการทำงานออกเป็น 2 ขั้นตอนวิธี คือ 1. ขั้นตอนแบบออฟไลน์ โดยใช้ขั้นตอนวิธี LCA-LDA (content based) probabilistic mixture generative ร่วมกับพิจารณาความชอบของผู้ใช้และความชอบของผู้เชี่ยวชาญภายในพื้นที่เหล่านั้นในการเลือกที่จะเยี่ยมชมสถานที่หรือเลือกที่จะเลือกทำกิจกรรมที่น่าสนใจ เนื่องจากความรู้ในพื้นที่ใหม่ที่ผู้ใช้พึ่งทำการเยี่ยมชมนั้น ผู้ใช้มีความรู้ในสถานที่ท่องเที่ยวที่น้อยจึงจำเป็นต้องทำการพิจารณาถึงผู้เชี่ยวชาญภายในพื้นที่ร่วมด้วย 2. ขั้นตอนแบบออนไลน์ ในขั้นตอนนี้ผู้ใช้จะทำการร้องขอพื้นที่ที่ผู้ใช้ต้องการเยี่ยมชมและระบบจะทำการสร้างรายการแนะนำจากขั้นตอนแบบออฟไลน์ร่วมกับการใช้ขั้นตอนวิธี Threshold Algorithm (TA) เพื่อให้การสร้างรายการแนะนำแบบออนไลน์สามารถใช้เวลาในการประมวลผลที่รวดเร็ว จากผลการทดลองขั้นตอนวิธี LCARS สามารถสร้างรายการแนะนำได้อย่างถูกต้องเมื่อเปรียบเทียบกับขั้นตอนวิธีการกรองร่วมแบบเดิมและใช้เวลาในการประมวลได้อย่างรวดเร็วอย่างมีประสิทธิภาพ

Zhang และคณะ (Zhang et al., 2015) ได้มีการนำเสนอขั้นตอนวิธีการสร้างรายการแนะนำสถานที่โดยเรียกว่า Opinion-based POI recommendation framework (ORec) ซึ่งใช้โมเดลการเรียนรู้แบบมีผู้สอนโดยการวิเคราะห์แง่อารมณ์ของความคิดเห็นที่ผู้ใช้ให้กับสถานที่นั้น ๆ เช่น ด้านการบริการ ราคา บรรยากาศ รสชาติ เป็นต้น รวมถึงความสัมพันธ์ระหว่างผู้ใช้และระยะทางระหว่างผู้ใช้กับสถานที่ ๆ ต้องการแนะนำ ข้อดีของวิธีการนี้คือสามารถสร้างรายการแนะนำสถานที่

ใหม่ที่ผู้ใช้มีความชื่นชอบและสามารถช่วยให้เจ้าของธุรกิจสร้างโฆษณาที่ดึงดูดผู้ใช้ที่มีความสนใจในธุรกิจนั้น ๆ เพื่อเพิ่มยอดขายได้อีกด้วย จากผลการทดลองแสดงให้เห็นว่าขั้นตอนวิธี ORec สามารถสร้างรายการแนะนำได้อย่างถูกต้องเมื่อเปรียบเทียบกับขั้นตอนวิธีการวิเคราะห์แง่อารมณ์แบบดั้งเดิม

Zhang และ Chow (Zhang & Chow, 2016) ได้นำเสนอขั้นตอนวิธีการสร้างรายการแนะนำสถานที่ ๆ เรียกว่า Temporal Influence Correlations for time-aware location Recommendations (TICRec) โดยการคำนึงถึงช่วงเวลาในการเยี่ยมชมสถานที่ของผู้ใช้และใช้เทคนิค Kernel density estimation (KDE) ในการแก้ไขปัญหาความเบาบางของข้อมูลเนื่องจากวิธีการโดยส่วนมากจะมีการแบ่งเวลาออกเป็นช่วง ๆ ซึ่งแสดงถึงการแบ่งส่วนข้อมูลออกเป็นส่วนย่อย ๆ แต่เทคนิค KDE ได้ใช้ข้อมูลเวลาแบบต่อเนื่องไม่มีการแบ่งข้อมูลซึ่งส่งผลให้ไม่ประสบปัญหาความเบาบางของข้อมูล อีกทั้งยังมีการคำนึงถึงความสัมพันธ์ที่ผู้ใช้มีการเช็คอินร่วมกันในสถานที่เดียวกันแต่คนละช่วงเวลาและผู้ใช้คนเดียวกันมีการเยี่ยมชมสถานที่ ๆ แตกต่างกันในแต่ละช่วงเวลาด้วยการใช้เทคนิควิธีการกรองร่วมร่วมกับพิจารณาผู้ใช้เป็นหลักและพิจารณาสถานที่เป็นหลัก จากผลการทดลองวิธีการที่นำเสนอ TICRec สามารถสร้างรายการแนะนำได้อย่างถูกต้องและไม่ประสบปัญหาความเบาบางของข้อมูล

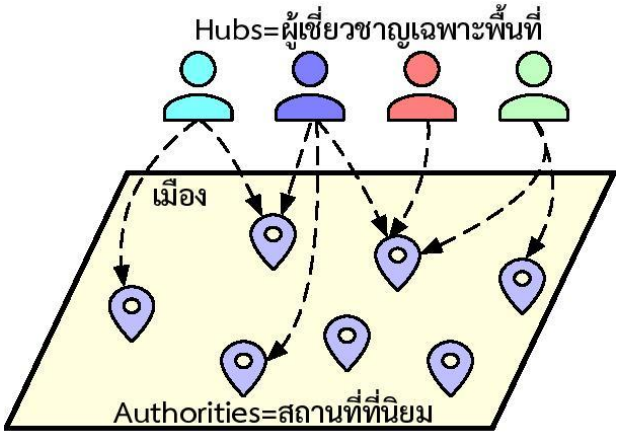
Si และคณะ (Si et al., 2017) นำเสนอขั้นตอนวิธีการสร้างรายการแนะนำสถานที่โดยเรียกว่า CTF-ARA ซึ่งทำการแบ่งประเภทผู้ใช้งานออกเป็นผู้ใช้ซึ่งมีการเช็คอินในสถานที่ต่าง ๆ บ่อยครั้ง (Active user) และผู้ใช้ซึ่งมีการเช็คอินในสถานที่ต่าง ๆ เพียงครั้งคราว (Inactive user) ด้วยเทคนิคการจัดกลุ่ม K-means โดยผู้ใช้ที่เป็นประเภทมีการเช็คอินบ่อยครั้งจะมีการสร้างรายการแนะนำโดยการคำนึงถึงช่วงเวลาในการเช็คอินของผู้ใช้และแก้ไขปัญหาความเบาบางของข้อมูลด้วยการหาความคล้ายของสถานที่ใกล้เคียงในชั่วโมงต่าง ๆ กับชั่วโมงในปัจจุบันที่ต้องการสร้างรายการแนะนำ จากผลการทดลองแสดงให้เห็นว่าขั้นตอนวิธีที่นำเสนอนี้สามารถสร้างรายการแนะนำได้อย่างถูกต้องและแก้ไขปัญหาความเบาบางของข้อมูลได้

Kefalas และคณะ (Kefalas et al., 2017) นำเสนอขั้นตอนวิธีการสร้างรายการแนะนำผู้ใช้และสถานที่โดยใช้เทคนิคที่เรียกว่า Random Walk with Restart on Heterogeneous Spatio-Temporal algorithm (RWR-HST) ซึ่งเป็นวิธีการกรองร่วมแบบโมเดลในรูปแบบกราฟที่ซึ่งได้ทำการพิจารณาความสัมพันธ์ของการ เชคอินใน 2 ลักษณะคือ 1. การเชคอินของผู้ใช้กับสถานที่ซึ่งแสดงถึงความชอบของผู้ใช้ในระยะยาว 2. สถานที่กับเซสชันคือการเชคอินของผู้ใช้เฉพาะช่วงเวลาซึ่งแสดงถึงความชอบของผู้ใช้ในระยะสั้น จากผลการทองที่นำเสนอแสดงให้เห็นถึงวิธี RWR-HST สามารถสร้างรายการแนะนำได้อย่างถูกต้องกว่าวิธีการแบบเดิมซึ่งไม่มีการพิจารณาบริบททางด้านเวลาได้อย่างมีประสิทธิภาพ

ต่อมา Jiuxin (Jiuxin et al., 2018) นำเสนอขั้นตอนวิธีการแนะนำสถานที่โดยพิจารณารูปแบบการเช็คอินของผู้ใช้ที่ซึ่งสามารถแบ่งออกได้เป็น 3 ประเภทคือ สถานที่ที่คำนึงถึงเวลาที่ผู้ใช้เช็คอิน ความนิยมของสถานที่และความชอบส่วนตัวของผู้ใช้แล้วโมเดลจำแนกประเภทข้อมูลจะถูกนำมาใช้เพื่อทำนายสถานที่ถัดไปที่คาดว่าผู้ใช้จะทำการเยี่ยมชม

2.2 การสร้างรายการแนะนำสถานที่โดยพิจารณาผู้ใช้ภายในพื้นที่นั้น ๆ

งานวิจัยโดยส่วนมากทางด้านระบบแนะนำสถานที่มีความต้องการประวัติการเช็คอินของผู้ใช้ ซึ่งในกรณีของผู้ใช้ใหม่หรือนักท่องเที่ยวที่มีการเช็คอินน้อยครั้งหรือไม่มีประวัติการเช็คอินเลยจะส่งผลให้ความถูกต้องในการ สร้างรายการแนะนำลดลง ดังนั้นอีกแนวทางหนึ่งของงานวิจัยได้มีการศึกษาการใช้ประโยชน์จากผู้เชี่ยวชาญเฉพาะพื้นที่หรือผู้ใช้ภายในพื้นที่นั้น ๆ ในการช่วยค้นหาสถานที่ที่ได้รับความนิยมในพื้นที่นั้น ๆ ดังเช่น



ภาพที่ 2.2 ขั้นตอนวิธีการ Hyperlink-Induced Topic Search (HITS)

Zheng (Zheng et al., 2009) นำเสนอโมเดลที่เรียกว่า Tree-Based Hierarchical Graph (TBHG) สำหรับการสร้างรายการแนะนำสถานที่โดยพิจารณาการเช็คอินของผู้ใช้ในหลาย ๆ ขนาดของพื้นที่แล้วทำการค้นหาผู้เชี่ยวชาญเฉพาะพื้นที่เพื่อสร้างรายการแนะนำโดยการขั้นตอนวิธีการ HITS ดังภาพที่ 2.2 โดยขั้นตอนวิธี TBHG สามารถสร้างรายการแนะนำสถานที่ที่ถูกต้องกว่าระบบแนะนำสถานที่แบบดั้งเดิม

Bao (Bao et al., 2012) ได้นำเสนอขั้นตอนวิธีการสร้างรายการแนะนำสถานที่เฉพาะบุคคลโดยการค้นหาผู้เชี่ยวชาญเฉพาะพื้นที่ด้วยวิธี HITS ที่ซึ่งพิจารณาความถี่ของการเช็คอินของผู้ใช้และ

สถานที่ในแต่ละเมือง แล้วทำการสร้างรายการแนะนำสถานที่โดยการพิจารณาความคล้ายของหมวดหมู่ของการเช็คอินระหว่างผู้เชี่ยวชาญเฉพาะพื้นที่กับผู้ใช้เป้าหมายจากโครงสร้าง Weight Category Hierarchy (WCH) ที่ซึ่งวิธีการนี้สามารถสร้างรายการแนะนำให้กับนักท่องเที่ยวได้อย่างถูกต้อง

Long (Long, X. & Joshi, J., 2013) นำเสนอขั้นตอนวิธีการสร้างรายการแนะนำสถานที่ตามความนิยมเฉพาะพื้นที่โดยปรับปรุงขั้นตอนวิธี HITS ร่วมกับขั้นตอนวิธี Entropy ที่ซึ่งเพิ่มการพิจารณาบริบทของความหลากหลายในการเช็คอินสถานที่ของผู้ใช้ จากแนวคิดที่ว่าผู้ใช้ที่ไปในหลายสถานที่มักมีความรู้หรือมีความเชี่ยวชาญมากกว่าผู้ใช้ที่ไปเพียงสถานที่เดียว อีกทั้งวิธีการที่นำเสนอนี้มีการพิจารณาความชอบของเพื่อนของผู้ใช้เป้าหมายอีกด้วย โดยวิธีที่นำเสนอนี้สามารถสร้างรายการแนะนำที่มีความถูกต้องมากกว่าขั้นตอนวิธี HITS แบบดั้งเดิม

ในปี 2016 ขั้นตอนวิธี Personalized successive POI recommendation approach (Chen et al., 2016) ได้ถูกนำเสนอ ที่ซึ่งขั้นตอนวิธีนี้ได้พิจารณาพฤติกรรมการเช็คอินที่ต่อเนื่องกัน พฤติกรรมการเช็คอินภายในพื้นที่นั้น ๆ และความชอบของการเช็คอินสถานที่ภายในกลุ่มเพื่อน และได้ประยุกต์ใช้ขั้นตอนวิธี Distance-weighted HITS สำหรับค้นหาสถานที่ที่มีความน่าสนใจโดยพิจารณาถึงระยะทางหรือระแวกใกล้เคียงของที่อยู่ผู้ใช้ในปัจจุบันร่วมกับพิจารณาความชอบของการเช็คอินสถานที่ภายในกลุ่มเพื่อน โดยวิธีการที่นำเสนอนี้สามารถสร้างรายการแนะนำได้ถูกต้องมากกว่าวิธีการก่อนหน้า

Bagci และ Karagoz (Bagci & Karagoz, 2016) ได้นำเสนอขั้นตอนวิธีการสร้างรายการแนะนำสถานที่ ที่เรียกว่าขั้นตอนวิธี Random-walk ที่ซึ่งใช้พื้นฐานของโครงสร้างกราฟ โดยสามารถช่วยลดปัญหาความเบาบางของข้อมูลได้และเมื่อผู้ใช้ใหม่เข้ามาในระบบ โมเดลนี้ไม่จำเป็นต้องทำการฝึกฝนโมเดลใหม่ทั้งหมดซึ่งดีกว่าขั้นตอนวิธีการกรองร่วมแบบเดิม อีกทั้งยังมีการพิจารณาสถานที่ที่เป็นที่นิยม ความเป็นเพื่อนของผู้ใช้ และผู้เชี่ยวชาญเฉพาะพื้นที่ จากขั้นตอนวิธีการดังกล่าวสามารถให้ความถูกต้องได้ดีกว่าขั้นตอนวิธีการ Random-walk แบบเดิม

ต่อมา Ying (Ying et al., 2017) นำเสนอขั้นตอนวิธีการสร้างรายการแนะนำสถานที่โดยพัฒนาขั้นตอนวิธีการกรองข้อมูลร่วมที่เรียกว่า Context-aware tensor decomposition (CTD) ร่วมกับขั้นตอนวิธี HITS ที่ซึ่งพิจารณาความชอบของผู้ใช้ หมวดหมู่และช่วงเวลาของสถานที่ รวมถึงความชอบของเพื่อนของผู้ใช้เป้าหมาย โดยขั้นตอนวิธีการนี้สามารถแก้ไขปัญหาความเบาบางของข้อมูลและสามารถช่วยสร้างรายการแนะนำได้ถูกต้องมากกว่าวิธีการก่อนหน้า

ล่าสุดในปี 2019 ขั้นตอนวิธีการแนะนำรายการสถานที่สำหรับชอปปิงถูกนำเสนอ (Puspitaningrum et al., 2019) โดยการพิจารณาผู้เชี่ยวชาญเฉพาะพื้นที่และผู้เชี่ยวชาญในหมวดหมู่ของสถานที่ชอปปิงนั้น ๆ โดยแต่ละสถานที่สำหรับแนะนำให้กับผู้ใช้จะถูกเลือกโดยการ

วิเคราะห์เนื้อหาของข้อเสนอแนะซึ่งให้ผลลัพธ์ในแง่ดี จากนั้นขั้นตอนวิธี PageRank และ Lazy random walk จะถูกใช้ในการเลือกผู้เชี่ยวชาญเพื่อสร้างรายการแนะนำสถานที่ให้กับผู้ใช้ต่อไป โดยขั้นตอนวิธีนี้สามารถสร้างรายการแนะนำได้ถูกต้องมากกว่าวิธีการ PageRank และ Random walk แบบเดิม

จากงานวิจัยที่กล่าวมาในขั้นต้น หลาย ๆ งานวิจัยมีการประยุกต์ใช้การค้นหาผู้เชี่ยวชาญเฉพาะพื้นที่หรือผู้ใช้ในพื้นที่นั้นร่วมกับพิจารณาประวัติของผู้ใช้เพื่อใช้ในการสร้างรายการแนะนำให้มีความถูกต้องมากขึ้น แต่อย่างไรก็ตามการพิจารณาประวัติของผู้ใช้อาจมีความไม่เหมาะสมสำหรับผู้ใช้ใหม่ที่ซึ่งมีประวัติการเช็คอินน้อยครั้งหรือไม่เคยเช็คอินมาก่อน สำหรับในงานวิจัยของ Long และ Joshi (Long, X. & Joshi, 2013) ได้พิจารณาความหลากหลายของสถานที่โดยใช้วิธี Entropy แต่ว่าการใช้วิธีนี้มีความอ่อนไหวของข้อมูลสูงและไม่มีประสิทธิภาพในกรณีที่ถ้าผู้ใช้มีการไปเยี่ยมชมทุกสถานที่อย่างละ 1 ครั้งและอีกผู้ใช้มีการไปเยี่ยมชมเกือบทุกที่แต่ไปมากกว่า 1 ครั้ง จากการคำนวณด้วยวิธีการ Entropy นี้ ผู้ใช้คนที่สองจะมีค่าคะแนนความหลากหลายที่น้อยกว่าผู้ใช้คนแรกมาก แต่ในความเป็นจริงผู้ใช้คนที่สองอาจจะมีความเชี่ยวชาญที่มากกว่าหรือใกล้เคียงกับผู้ใช้คนแรก สำหรับขั้นตอนวิธีของ Long และ Joshi (Long, X. & Joshi, 2013) Bagci และ Karagoz (Bagci & Karagoz, 2016) Ying (Ying et al., 2017) Puspitaningrum (Puspitaningrum et al., 2019) ได้มีการพิจารณาความสัมพันธ์ของเพื่อนสำหรับการค้นหาสถานที่ที่น่าสนใจที่ซึ่งในบางครั้งความชอบของเพื่อนอาจมีความชอบที่ไม่คล้ายกัน แต่การพิจารณาผู้ใช้คนอื่นในระบบอาจจะมีความคล้ายกับผู้ใช้เป้าหมายมากกว่า สำหรับในงานของ Chen (Chen et al., 2016) ได้มีการพิจารณาระยะทางในการค้นหาสถานที่ที่น่าสนใจที่ซึ่งขั้นตอนวิธีนี้ใช้เวลาในการคำนวณที่นานในขั้นตอนการแนะนำแบบออนไลน์เพื่อค้นหาระยะห่างระหว่างที่อยู่ของผู้ใช้กับสถานที่ในระแวกใกล้เคียงเพื่อสร้างรายการแนะนำ สำหรับงานของ Puspitaningrum (Puspitaningrum et al., 2019) มีการพิจารณาความคิดเห็นแง่ดีจากผู้ใช้ในระบบเพื่อค้นหาสถานที่ที่น่าสนใจ ซึ่งในความคิดเห็นนั้นอาจจะมีความคิดเห็นปลอมแฝงอยู่ (Fake review) และต้องใช้โมเดลในการจำแนกความคิดเห็นเพื่อให้ได้ความถูกต้องที่สูง ดังแสดงปัจจัยที่ใช้ในการพิจารณาเพื่อค้นหาสถานที่ที่น่าสนใจ ดังตารางที่ 2.1 ดังนั้นในงานวิจัยที่นำเสนอจึงได้มีการสร้างรายการแนะนำสถานที่ที่น่าสนใจ โดยพิจารณาความถี่ในการเช็คอินของผู้ใช้ร่วมกับความชอบของผู้ใช้โดยคำนึงถึงความถี่ในการเช็คอินร่วมกับการกลับมาเยี่ยมชมของผู้ใช้ ของผู้ใช้ภายในพื้นที่ โดยปรับปรุงจากขั้นตอนวิธีการ Hyperlink-Induced Topic Search (HITS) จากขั้นตอนวิธีการที่นำเสนอโดยวิธีการที่เรานำเสนอนี้ไม่จำเป็นต้องใช้ประวัติของผู้ใช้และสามารถคำนวณเพื่อหาสถานที่ที่น่าสนใจในขั้นตอนออฟไลน์ทั้งหมดและในขั้นตอนออนไลน์จะทำการเรียงลำดับค่าคะแนนจากมากไปน้อยของสถานที่ในระแวกใกล้เคียงเท่านั้น

ตารางที่ 2.1 ปัจจัยที่ใช้ในการพิจารณาเพื่อค้นหาสถานที่ที่น่าสนใจ

งานวิจัย	คุณลักษณะของการพิจารณาสถานที่ที่มีความน่าสนใจและผู้เชี่ยวชาญเฉพาะพื้นที่						
	ความถี่	ความหลากหลาย	การกลับมาอีกครั้ง	ความคิดเห็น	ระยะทาง	เพื่อนของผู้ใช้	ประวัติผู้ใช้
Zheng et al., 2009	✓	–	–	–	–	–	✓
Bao et al., 2012	✓	–	–	–	–	–	✓
Long, X. & Joshi, 2013	✓	✓	–	–	–	✓	✓
Bagci & Karagoz, 2016	✓	–	–	–	–	✓	✓
Chen et al., 2016	✓	–	–	–	✓	–	✓
Ying et al., 2017	✓	✓	–	–	–	✓	✓
Puspitaningrum et al., 2019	–	–	–	✓	–	✓	–
ขั้นตอนวิธีการที่นำเสนอ	✓	–	✓	–	–	–	–

2.4 ฐานข้อมูลที่ใช้ในการทดลอง

ฐานข้อมูลที่ใช้ในการทดลองจะทำการจัดเก็บประวัติการเช็คอินของผู้ใช้จากเว็บไซต์ Foursquare (Yang et al., 2015) ประกอบด้วย 1. รหัสผู้ใช้ 2. รหัสสถานที่ 3. รหัสหมวดหมู่สถานที่ 4. หมวดหมู่สถานที่ 5. ละติจูด 6. ลองจิจูด 7. เวลาในการเช็คอิน 8. เวลา UTC โดยจัดเก็บในช่วง 12 เมษายน 2012 ถึง 16 กุมภาพันธ์ 2013 (ประมาณ 10 เดือน) แสดงตัวอย่างของฐานข้อมูล Foursquare ที่ใช้ในการทดลอง ดังภาพที่ 2.3

รหัสผู้ใช้	รหัสสถานที่	รหัสหมวดหมู่สถานที่	หมวดหมู่สถานที่	ละติจูด	ลองจิจูด	เวลาในการเช็คอิน	เวลา UTC
1386	4b64f0e3f964a5203cdb2ae3	4bf58dd8d48988d16e941735	Fast Food Restaurant	35.577	19.660	540	Tue Apr 03 21:33:17 +0000 2012
2243	4f534895e4b006151584808a	4bf58dd8d48988d1c4941735	Restaurant	35.627	39.714	40	Tue Apr 03 21:50:15 +0000 2012
1064	4b6fe136f964a520b4fe2ce3	4bf58dd8d48988d1e0931735	Coffee Shop	35.689	39.700	40	Tue Apr 03 22:13:21 +0000 2012
1248	4b72c558f964a520d2872de3	4bf58dd8d48988d111941735	Japanese Restaurant	35.652	39.544	540	Tue Apr 03 22:23:00 +0000 2012
1540	4b5829f4f964a520a44c28e3	4bf58dd8d48988d1e0931735	Coffee Shop	35.691	39.703	40	Tue Apr 03 22:23:18 +0000 2012
1964	4b5e4928f964a520c68729e3	4bf58dd8d48988d16e941735	Fast Food Restaurant	35.646	39.745	540	Tue Apr 03 22:25:30 +0000 2012
2243	4f534895e4b006151584808a	4bf58dd8d48988d1c4941735	Restaurant	35.627	39.7143	40	Tue Apr 03 22:26:33 +0000 2012
9655	4cde2fce41ed224b1e97d73c	4d4ae6fc7a7b7dea34424761	Fried Chicken Joint	35.789	39.661	540	Tue Apr 03 22:27:30 +0000 2012
1100	4b5a3b6cf964a52026b628e3	4bf58dd8d48988d16e941735	Fast Food Restaurant	35.574	139.658	540	Tue Apr 03 22:27:54 +0000 2012
1320	4df7423e483b96f73159a256	4bf58dd8d48988d111941735	Japanese Restaurant	35.598	39.6669	540	Tue Apr 03 22:29:18 +0000 2012

ภาพที่ 2.3 ตัวอย่างของฐานข้อมูล Foursquare ที่ใช้ในการทดลอง

ในการทดลองนี้ได้คัดเลือกใช้เฉพาะหมวดหมู่อาหารและทำการทดลองใน 5 พื้นที่ของเมืองโตเกียวคือ โดยารแปลงละติจูดและลองจิจูดไปยังพื้นที่ต่าง ๆ ด้วย Geocoding API[5] ที่ซึ่งคุณลักษณะของเมืองต่าง ๆ และในงานวิจัยที่นำเสนอได้มีการพิจารณาผู้ใช้ที่มีการเช็คอินมากกว่า 1 สถานที่และสถานที่ต้องถูกเช็คอินโดนผู้ใช้มากกว่า 1 ผู้ใช้ เพื่อป้องกันสถานที่ที่ถูกสร้างขึ้นเฉพาะตนเองและลดความเบางบางของข้อมูลแสดงคุณลักษณะของเมืองต่าง ๆ ที่ถูกใช้ในการทดลอง ดังตารางที่ 2.2

ตารางที่ 2.2 คุณลักษณะของเมืองต่าง ๆ ที่ถูกใช้ในการทดลอง

พื้นที่	**ผู้ใช้**	**สถานที่**	**จำนวนเช็คอินทั้งหมด**	**จำนวนเช็คอินต่อสถานที่**	**จำนวนเช็คอินต่อผู้ใช้**
Chiyoda	958	827	7283	8.81	7.60
Shibuya	581	719	4107	5.71	7.10
Minato	519	719	3836	5.34	7.39
Shinjuku	571	670	3905	6.84	5.83
Chuo	243	325	1645	5.06	6.77

[5] https://developers.google.com/maps/documentation/geocoding/

บทที่ 3
วิธีการที่นำเสนอ

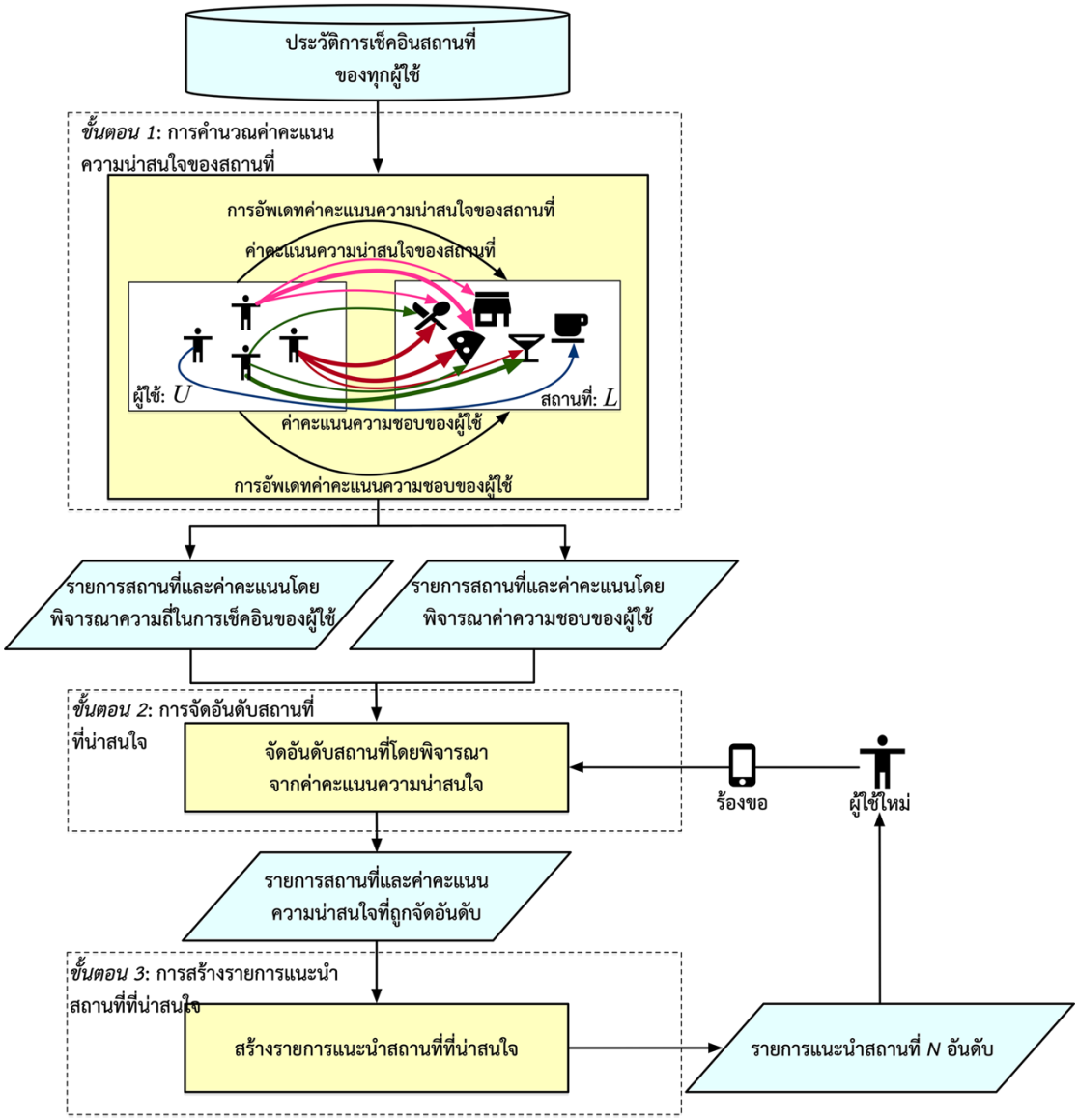
ในส่วนนี้จะกล่าวถึงนิยามพื้นฐานของปัญหาการสร้างรายการแนะนำสถานที่และรายละเอียดของขั้นตอนวิธีการที่นำเสนอสำหรับสร้างรายการแนะนำสถานที่ที่เรียกว่า ระบบแนะนำสถานที่ที่น่าสนใจมากที่สุดเอ็นอันดับ (N-most Interesting location-based recommender system , NILR) ที่ซึ่งวิธีการที่นำเสนอนี้จะพิจารณาสถานที่ในหมวดหมู่อาหาร เช่น ร้านอาหาร ร้านกาแฟ คาเฟ่ และ ร้านเบเกอรี่ เป็นต้น และให้ค่าคะแนนกับความน่าใจของสถานที่โดยพิจารณา 2 ปัจจัย คือ 1. ความถี่ในการเช็คอินของผู้ใช้ และ 2. ความชอบของผู้ใช้โดยคำนึงถึงความถี่ในการเช็คอินร่วมกับการกลับมาเยี่ยมชมของผู้ใช้ จากนั้นจะทำการจัดอันดับโดยพิจารณาค่าคะแนนที่มากที่สุดของทั้งสองปัจจัยเพื่อสร้างเป็นรายการแนะนำสถานที่ที่น่าสนใจ N อันดับให้กับผู้ใช้ใหม่โดยคำนึงถึงที่อยู่ปัจจุบันของผู้ใช้และสถานที่ในรายการแนะนำจะเป็นสถานที่ในระแวกใกล้เคียงกับที่อยู่ปัจจุบันของผู้ใช้ และสุดท้ายตัวอย่างการสร้างรายการแนะนำสถานที่ที่น่าสนใจด้วยวิธีที่นำเสนอจะถูกอธิบายในส่วนนี้

3.1 นิยามพื้นฐาน

กำหนดให้ C เป็นพื้นที่เป้าหมายสำหรับการสร้างรายการแนะนำสถานที่ ซึ่งประกอบด้วยเซทของร้านอาหาร $L = \{l_1, l_2, \dots, l_p\}$ จำนวน p ร้านอาหารที่ซึ่งตั้งอยู่ในพื้นที่เป้าหมาย C ที่ซึ่ง l_j เป็นสมาชิกของ L ($l_j \in L$) โดยแต่ละสถานที่ประกอบด้วยลองจิจูดและละติจูดสำหรับกำหนดที่อยู่ของสถานที่นั้น ๆ ($< lat_{l_j}, lon_{l_j} >$) ดังนั้นเซทของร้านอาหารสามารถถูกเขียนใหม่ได้เป็น $L = \{< l_1, lat_{l_1}, lon_{l_1} >, < l_2, lat_{l_2}, lon_{l_2} >, \dots, < l_q, lat_{l_q}, lon_{l_m} >\}$ จากนั้นกำหนดให้เซทของผู้ใช้ $U = \{u_1, u_2, \dots, u_q\}$ จำนวน q ผู้ใช้ ซึ่งผู้ใช้ต้องมีการเช็คอินอย่างน้อย 1 ร้านอาหารในเซท L โดยขั้นตอนวิธีการที่นำเสนอสำหรับสร้างรายการแนะนำร้านอาหารจะถูกคำนวณในรูปแบบของเมทริกซ์ $A \in \mathbb{R}^{p \times q}$ ที่ซึ่งแต่ละแถวในเมทริกซ์แสดงถึงแต่ละผู้ใช้ $u_k \in U$ และแต่ละคอลัมภ์แสดงถึงแต่ละสถานที่ $l_j \in L$ สำหรับแต่ละข้อมูลภายในเมทริกซ์คือจำนวนการเช็คอินของผู้ใช้ u_k มีการเช็คอินในร้านอาหาร l_j โดยสามารถกำหนดได้เป็น f_{u_k,l_j} ดังนั้นปัญหาของการแนะนำสถานที่ในการค้นหา N อันดับที่น่าสนใจมากที่สุดของสถานที่สำหรับแนะนำให้กับผู้ใช้ใหม่ nu ที่ซึ่ง $nu \notin U$ ผู้ซึ่งต้องการรายการแนะนำสถานที่ในระแวกใกล้เคียงกับที่อยู่ปัจจุบันของผู้ใช้ สามารถกำหนดได้เป็น $L^{nu} = \{l_1, l_2, \dots, l_n\}$ ทั้งหมด n รายการที่ $l_j \in L^{nu}$ ซึ่งเป็นสถานที่ที่มีความสนใจมากที่สุดที่ผู้ใช้ใหม่ nu ควรจะได้รับคำแนะนำ

3.2 ระบบแนะนำสถานที่ที่น่าสนใจมากที่สุดเอ็นอันดับ (N-most Interesting location-based recommender system, NILR)

วิธีการที่นำเสนอต้องการประวัติการเช็คอินของผู้ใช้เป็นข้อมูลนำเข้า จากนั้นรายการแนะนำสถานที่ที่สนใจ N อันดับโดยการพิจารณาประวัติการเช็คอินของผู้ใช้ในระบบ โดยวิธีการที่นำเสนอประกอบด้วย 3 ขั้นตอนการทำงานหลัก คือ 1. การคำนวณค่าคะแนนความน่าสนใจของสถานที่ 2. การจัดอันดับสถานที่โดยพิจารณาค่าคะแนนความน่าสนใจ 3. การสร้างรายการแนะนำสถานที่ N อันดับ แสดงวิธีการทำงานดังภาพที่ 3.1



ภาพที่ 3.1 ขั้นตอนวิธีการสร้างรายการแนะนำสถานที่ NILR

3.2.1 การคำนวณค่าคะแนนความน่าสนใจของสถานที่

ในการคำนวณค่าคะแนนความน่าสนใจของสถานที่ในพื้นที่ C ทุกประวัติการเช็คอินของผู้ใช้ในพื้นที่นี้จะถูกพิจารณา โดยค่าคะแนนความน่าสนใจของแต่ละสถานที่ l_j จะถูกคำนวณแบบวนซ้ำ ที่ซึ่งค่าคะแนนความน่าสนใจของสถานที่จะถูกคำนวณโดยการพิจารณาผลรวมความถี่ในการเช็คอินของผู้ใช้ร่วมกับค่าคะแนนของผู้ใช้ที่ซึ่งพิจารณาจากค่าคะแนนความน่าสนใจของสถานที่ l_j และความถี่ที่ผู้ใช้เช็คอินในสถานที่ l_j ที่ซึ่งในขั้นตอนวิธีการที่นำเสนอนี้ได้ใช้ขั้นตอนวิธี HITS ในการคำนวณค่าคะแนนความน่าสนใจของสถานที่ซึ่งมีรายละเอียดวิธีการทำงาน ดังนี้

นิยามที่ 1 ค่าคะแนนความน่าสนใจของสถานที่โดยพิจารณาความถี่ในการเช็คอินของผู้ใช้

กำหนดให้ $L = \{l_1, l_2, \dots, l_p\}$ เป็นเซทรายการสถานที่ซึ่งมีค่าคะแนนความน่าสนใจของสถานที่โดยพิจารณาความถี่ในการเช็คอินของผู้ใช้ ประกอบด้วย $ISF^L = \{isf(l_1), isf(l_2), \dots, isf(l_p)\}$ และให้ผู้ใช้ที่มีการเช็คอินในสถานที่ l_j ใดๆสามารถกำหนดเป็นเซทของผู้ใช้ $U = \{u_1, u_2, \dots, u_q\}$ ที่ซึ่งแต่ละผู้ใช้มีค่าคะแนนประกอบด้วย $ISF^U = \{isf(u_1), isf(u_2), \dots, isf(u_q)\}$ ดังนั้นค่าคะแนนของสถานที่ l_j ใดๆ ($isf(l_j)$) โดยพิจารณาความถี่ในการเช็คอินของผู้ใช้ สามารถคำนวณได้จากผลรวมของผลคูณระหว่างความถี่ในการเช็คอินของผู้ใช้ u_k ต่อสถานที่ l_j (f_{u_k,l_j}) กับค่าคะแนนของผู้ใช้ u_k ($isf(u_k)$) ในรอบแรกค่าคะแนนของผู้ใช้จะถูกกำหนดเป็น 1 และรอบต่อไปจะถูกคำนวณจากผลรวมของผลคูณระหว่างความถี่ในการเช็คอินของผู้ใช้ u_k ต่อสถานที่ l_j (f_{u_k,l_j}) กับค่าคะแนนของสถานที่ l_j ($isf(l_j)$) เมื่อค่าคะแนนของผู้ใช้ u_k ถูกปรับใหม่ ทำให้ต้องเริ่มการคำนวณค่าคะแนนของสถานที่ l_j ใหม่ จากขั้นตอนวิธีนี้มีการคำนวณแบบวนซ้ำจนกว่าค่าคะแนนคงที่ โดยแต่ละรอบขั้นตอนวิธี *L2*-normalization จะถูกนำมาใช้ในการปรับค่าคะแนน ISF^L และ ISF^U ให้อยู่ในช่วง 0 ถึง 1 แสดงขั้นตอนวิธีการคำนวณค่าคะแนนความน่าสนใจของสถานที่โดยพิจารณาความถี่ในการเช็คอินของผู้ใช้ดังภาพที่ 3.2

Algorithm 1 HITS

Input: matrix $A_{q\times p}$, and x steps
Output: set of location scores ISF^L, and set of users ISF^U

1: $ISF^L = [1_{l_1}, 1_{l_2}, ...,1_{l_p}]$
2: $ISF^U = [1_{u_1}, 1_{u_2}, ...,1_{u_q}]$
3: **for each** x steps **do**
4: $ISF^L = A^T \cdot ISF^U$
5: $ISF^U = A \cdot ISF^L$
6: normalize ISF^L and ISF^U
7: **end for**
8: **return** ISF^L and ISF^U

ภาพที่ 3.2 ขั้นตอนวิธี HITS สำหรับคำนวณค่าคะแนนความน่าสนใจของสถานที่โดยพิจารณาความถี่ในการเช็คอินของผู้ใช้

จากภาพที่ 3.2 เริ่มจากการกำหนดค่าคะแนนของสถานที่และผู้ใช้ให้มีค่าคะแนนเท่ากับ 1 (บรรทัดที่ 1-2) จากนั้นแต่ละรอบการคำนวณ ค่าคะแนนสถานที่จะถูกคำนวณก่อน (บรรทัดที่ 4) และจากค่าคะแนนของสถานที่จะถูกนำไปคำนวณค่าคะแนนของผู้ใช้ (บรรทัดที่ 5) จากนั้นเซทค่าคะแนนสถานที่และผู้ใช้จะถูกปรับโดยใช้ขั้นตอนวิธี *L2*-normalization (บรรทัดที่ 6) โดยการคำนวณดังกล่าวจะถูกทำซ้ำตามจำนวน x รอบที่ได้กำหนดไว้ในขั้นแรก

นิยามที่ 2 ค่าคะแนนความน่าสนใจของสถานที่โดยพิจารณาค่าความชอบของผู้ใช้

กำหนดให้ $L = \{l_1, l_2, \dots, l_p\}$ เป็นเซทรายการสถานที่ซึ่งมีค่าคะแนนความน่าสนใจของสถานที่โดยพิจารณาค่าความชอบของผู้ใช้ ประกอบด้วย $ISP^L = \{isp(l_1), isp(l_2), \dots, isp(l_p)\}$ และให้ผู้ใช้ที่มีการเช็คอินในสถานที่ l_j ใด ๆ สามารถกำหนดเป็นเซทของผู้ใช้ $U = \{u_1, u_2, \dots, u_q\}$ ที่ซึ่งแต่ละผู้ใช้มีค่าคะแนนประกอบด้วย $ISP^U = \{isp(u_1), isp(u_2), \dots, isp(u_q)\}$ กำหนดให้ค่าความชอบของผู้ใช้ u_k ใด ๆ ($pref_{u_k}$) สามารถคำนวณจากจำนวนสถานที่ที่ผู้ใช้ u_j มีการเช็คอินมากกว่า 1 ครั้ง นั่นคือ $pref_{u_k} = \sum_{j,f_{u_k,l_j}>1}^{p} 1$ และค่าความชอบของสถานที่ l_jใด ๆ ($pref_{l_j}$) สามารถคำนวณจากจำนวนผู้ใช้ซึ่งมีการเช็คอินที่สถานที่ l_j มากกว่า 1 ครั้ง นั่นคือ $pref_{l_j} = \sum_{k,f_{u_k,l_j}>1}^{q} 1$ สำหรับค่าคะแนนความน่าสนใจของสถานที่โดยพิจารณาค่าความชอบของผู้ใช้ l_j ($isp(l_j)$) สามารถคำนวณได้โดยผลรวมของผลคูณระหว่างความถี่ในการเช็คอินของผู้ใช้ u_k ต่อสถานที่ l_j (f_{u_k,l_j}) กับค่าความชอบของผู้ใช้ u_k ($pref_{u_k}$) และค่าคะแนนของผู้ใช้ u_k ($isp(u_k)$) ในรอบแรกค่าคะแนน

ของผู้ใช้จะถูกกำหนดเป็น 1 และรอบต่อไปจะถูกคำนวณจากผลรวมของผลคูณระหว่างความถี่ในการเช็คอินของผู้ใช้ u_k ต่อสถานที่ l_j (f_{u_k,l_j}) กับค่าความชอบของสถานที่ l_j ($pref_{l_j}$) และค่าคะแนนของสถานที่ l_j ($isp(l_j)$) จากขั้นตอนวิธีนี้มีการคำนวณแบบวนซ้ำจนกว่าค่าคะแนนคงที่ โดยแต่ละรอบขั้นตอนวิธี *L2*-normalization จะถูกนำมาใช้ในการปรับค่าคะแนน ISP^L และ ISP^U ให้อยู่ในช่วง 0 ถึง 1 แสดงขั้นตอนวิธีการคำนวณค่าคะแนนความน่าสนใจของสถานที่โดยพิจารณาความถี่ในการเช็คอินของผู้ใช้ แสดงขั้นตอนวิธีการคำนวณค่าคะแนนความน่าสนใจของสถานที่โดยพิจารณาค่าความชอบของผู้ใช้ที่เรียกว่าขั้นตอนวิธี ReturnHITS ดังภาพที่ 3.3

Algorithm 2 ReturnHITS

Input: matrix $A_{q\times p}$, and x steps
Output: set of location scores ISP^L, and set of users ISP^U

1: $pref_{l_j}$ = preference of location l_j
2: $pref_{u_k}$ = preference of user u_k
3: $AL_{u_.,l_j} = f_{u_.,l_j} \times pref_{l_j}$
4: $AU_{u_k,l_.} = f_{u_k,l_.} \times pref_{u_k}$
5: $ISP^L = [1_{l_1}, 1_{l_2}, ..., 1_{l_p}]$
6: $ISP^U = [1_{u_1}, 1_{u_2}, ..., 1_{u_q}]$
7: **for each** x steps **do**
8: $\quad ISP^L = AU^T \cdot ISP^U$
9: $\quad ISP^U = AL \cdot ISP^L$
10: $\quad$ normalize ISP^L and ISP^U
11: **end for**
12: **return** ISP^L and ISP^U

ภาพที่ 3.3 ขั้นตอนวิธี ReturnHITS สำหรับคำนวณค่าคะแนนความน่าสนใจของสถานที่โดยพิจารณาค่าความชอบของผู้ใช้

จากภาพที่ 3.3 ขั้นตอนแรกค่าคะแนนความชอบของทุกสถานที่และทุกผู้ใช้จะถูกคำนวณ (บรรทัดที่ 1-2) จากนั้นเมทริกซ์ค่าความชอบของทุกสถานที่และทุกผู้ใช้จะถูกสร้างจากผลคูณของความถี่ในการเช็คอินกับค่าความชอบของสถานที่และความถี่ในการเช็คอินและค่าความชอบของผู้ใช้ตามลำดับ (บรรทัดที่ 3-4) แล้วค่าคะแนนความน่าสนใจของสถานที่และคะแนนความน่าสนใจของผู้ใช้จะถูกกำหนดให้มีค่าเริ่มต้นเป็น 1 (บรรทัดที่ 5-6) และค่าคะแนนความน่าสนใจของสถานที่จะถูกคำนวณ (บรรทัดที่ 8) จากนั้นค่าคะแนนความน่าสนใจของผู้ใช้จะถูกคำนวณ (บรรทัดที่ 9) เมื่อคะแนนความน่าสนใจของสถานที่และผู้ใช้ถูกคำนวณแล้วจะถูกปรับลดค่าให้อยู่ในช่วง 0 ถึง 1 โดยใช้

ขั้นตอนวิธี *L2*-normalization แล้วขั้นตอนวิธีจากบรรทัดที่ 8-10 จะถูกทำซ้ำตามจำนวน x รอบที่ได้กำหนดไว้ในขั้นแรก

เมื่อการคำนวณแบบวนซ้ำจนค่าคะแนนของทั้งทุกสถานที่และทุกผู้ใช้ไม่เปลี่ยนแปลง เราจะได้ค่าคะแนนของทั้งผู้ใช้และสถานที่โดยการพิจารณาความถี่ในการเช็คอินและค่าความชอบทั้งผู้ใช้และสถานที่ ดังนี้

$$ISF^L = \{isf(l_1), isf(l_2), \dots, isf(l_p)\},$$
$$ISF^U = \{isf(u_1), isf(u_2), \dots, isf(u_q)\},$$
$$ISP^L = \{isp(l_1), isp(l_2), \dots, isp(l_p)\},$$
$$ISP^U = \{isp(u_1), isp(u_2), \dots, isp(u_q)\}$$

3.2.2 การจัดอันดับสถานที่ที่น่าสนใจ

ในการสร้างรายการแนะนำสถานที่สำหรับแนะนำให้ผู้ใช้ใหม่ ค่าคะแนนความน่าสนใจของสถานที่โดยพิจารณาความถี่ในการเช็คอินของผู้ใช้ (ISF^L) และค่าคะแนนความน่าสนใจของสถานที่โดยพิจารณาค่าความชอบของผู้ใช้ (ISP^L) จะถูกพิจารณาโดยในกระบวนการนี้จะเป็นการพิจารณาค่าคะแนนที่มากที่สุดของ ISF^L กำหนดเป็น max^{freq} และค่าคะแนนที่มากที่สุดของ ISP^L กำหนดเป็น max^{pref} แบบวนซ้ำ ซึ่งจะเป็นการระบุถึงเซทของสถานที่ซึ่งมีค่าคะแนนมากที่สุดทั้งทางด้านความถี่ในการเช็คอินและความชอบของผู้ใช้ สามารถกำหนดเป็น L^{freq} และ L^{pref} ตามลำดับ จากนั้นทั้งสองค่าคะแนนจะถูกเปรียบเทียบสำหรับการพิจารณาอันดับของสถานที่ในรายการแนะนำ ถ้าทั้งสองค่าคะแนนมีค่าเท่ากัน ($max^{freq} = max^{pref}$) แล้วทั้งสองเซทของสถานที่จะถูกจัดอันดับเป็นอันดับเดียวกัน นอกจากนี้ถ้าค่าคะแนนที่มากที่สุดของสถานที่ทางด้านความถี่ในการเช็คอินมีค่าที่มากกว่าค่าคะแนนที่มากที่สุดของสถานที่ทางด้านความชอบของผู้ใช้ ($max^{freq} > max^{pref}$) แล้วทุกสถานที่ของเซททางด้านความถี่จะถูกจัดเป็นอันดับที่สูงกว่าทุกสถานที่ของเซททางด้านความชอบ ถ้าค่าคะแนนที่มากที่สุดของสถานที่ทางด้านความชอบมีค่าที่มากกว่าค่าคะแนนที่มากที่สุดของสถานที่ทางด้านความถี่ในการเช็คอิน ($max^{pref} > max^{freq}$) แล้วทุกสถานที่ของเซททางด้านความชอบจะถูกจัดเป็นอันดับที่สูงกว่าทุกสถานที่ของเซททางด้านความถี่ในการเช็คอิน เมื่อพิจารณาการจัดอันดับของทั้งสองเซทสถานที่ แล้วเซทของสถานที่จะถูกนำออกจากเซทของสถานที่ ISF^L และ ISP^L จากนั้นขั้นตอนวิธีการจัดอันดับจะถูกคำนวณต่อไปจนกระทั่งทุกสถานที่จะถูกจัดอันดับจนครบ แล้วทำการจัดเก็บในรายการแนะนำสถานที่สุดท้าย (IL) โดยรายละเอียดขั้นตอนวิธีการจัดอันดับรายการแนะนำสถานที่ แสดงดังภาพที่ 3.4

Algorithm 3 Ranking

Input: set of locations L, set of location scores ISF^L and ISP^L
Output: set of recommending list IL

1: $IL = \emptyset, rank = 1$
2: **while** $L \neq \emptyset$ **do**
3: $max^{freq} = max(ISF^L)$
4: $max^{pref} = max(ISP^L)$
5: $L^{freq} = \{l_j \in L | isf(l_j) = max^{freq}\}$
6: $L^{pref} = \{l_k \in L | isp(l_k) = max^{pref}\}$
7: **if** $max^{freq} = max^{pref}$ **then**
8: $IL = IL \cup \{< l_j, rank > | l_j \in L^{freq} \cup L^{pref}\}$
9: **else if** $max^{freq} > max^{pref}$ **then**
10: $IL = IL \cup \{< l_j, rank > | l_j \in L^{freq}\}$
11: $rank{+}{+}$
12: $IL = IL \cup \{< l_k, rank > | l_k \in L^{pref}\}$
13: **else**
14: $IL = IL \cup \{< l_k, rank > | l_k \in L^{pref}\}$
15: $rank{+}{+}$
16: $IL = IL \cup \{< l_j, rank > | l_j \in L^{freq}\}$
17: $ISF^L = ISF^L - \{isf(l_j) | l_j \in L^{freq}\}$
18: $ISP^L = ISP^L - \{isp(l_k) | l_k \in L^{pref}\}$
19: $L = L - (L^{freq} \cup L^{pref})$
20: $rank{+}{+}$
21: **return** IL

ภาพที่ 3.4 รายละเอียดขั้นตอนวิธีการจัดอันดับรายการแนะนำสถานที่

จากภาพที่ 3.4 ในขั้นแรกค่าคะแนนที่มากที่สุด (max^{freq} และ max^{pref}) จากเซทรายการค่าคะแนนความน่าสนใจของสถานที่โดยพิจารณาความถี่ในการเช็คอินของผู้ใช้ (ISF^L) และค่าคะแนนความน่าสนใจของสถานที่โดยพิจารณาค่าความชอบของผู้ใช้ (ISP^L) จะถูกค้นหาร่วมกับเซทรายการของค่าคะแนนที่มากที่สุด (L^{freq} และ L^{pref}) ดังกล่าว (บรรทัดที่ 3-6) จากนั้นค่าคะแนนที่มากที่สุดของทั้งสองเซทรายการจะถูกเปรียบเทียบกันใน 3 กรณีที่กล่าวมาขั้นต้น (บรรทัดที่ 7-16) แล้วค่าคะแนนของเซทรายการสถานที่ ISF^L ที่ถูกพิจารณาแล้วจะถูกลบออกจากการพิจารณา ISF^L (บรรทัดที่ 17) และค่าคะแนนของเซทรายการสถานที่ ISP^L ที่ถูกพิจารณาแล้วจะถูกลบออกจากการพิจารณา ISP^L (บรรทัดที่ 18) และเซทรายการสถานที่ของค่าคะแนนดังกล่าวจะถูกลบออกจากการพิจารณาของเซทรายการสถานที่ L สุดท้ายจะทำการเลื่อนอันดับของเซทรายการผลลัพธ์ IL (บรรทัดที่ 20) โดยขั้นตอนวิธีนี้จะถูกคำนวณแบบวนซ้ำจนกระทั่งเซทรายการสถานที่เป็นเซทว่างหรือไม่มีเซทรายการให้พิจารณา

3.2.3 การสร้างรายการแนะนำสถานที่ที่น่าสนใจมากที่สุด N อันดับ

ในการแนะนำสถานที่ให้กับผู้ใช้ใหม่ (nu) นั้น ระบบแนะนำสถานที่จำเป็นต้องมีที่อยู่ปัจจุบันของผู้ใช้ ($loc(nu)$) เพื่อค้นหาสถานที่ที่น่าสนใจในระแวกใกล้เคียง แล้วขั้นตอนวิธีการที่นำเสนอจะทำการเลือกรายการสถานที่ที่น่าสนใจมากที่สุด N อันดับจากรายการแนะนำสถานที่ IL โดยแต่ละสถานที่ที่เลือกจะต้องอยู่ในระยะที่ไม่ไกลกว่าอยู่ปัจจุบันของผู้ใช้ d กิโลเมตรและมีอันดับที่สูงที่สุด N อันดับ จากนั้นรายการแนะนำสถานที่ที่น่าสนใจ L^{nu}จะถูกแนะนำให้กับผู้ใช้ nu แสดงรายละเอียดขั้นตอนวิธีการสร้างรายการแนะนำสถานที่ที่น่าสนใจมากที่สุด N อันดับ ดังภาพที่ 3.5

Algorithm 4 Top-N Generation

Input: current location of user nu $loc(nu)$, distance d, the number of recommendation N, set of recommending list IL, set of locations L
Output: An order list of N-most interesting locations, L^{nu}

1: $L^{nu} = \emptyset, rank = 0$
2: **for** each location l_j in IL and $rank < N$ **do**
3: **if** diff$(loc(nu), loc(l_j)) \leq d$ **then**
4: $L^{nu} = L^{nu} \cup \{l_k | l_k \in IL, rank_{l_k} = rank_{l_j}\}$
5: $rank{+}{+}$
6: **return** L^{nu}

ภาพที่ 3.5 แสดงรายละเอียดขั้นตอนวิธีการสร้างรายการแนะนำสถานที่ที่น่าสนใจมากที่สุด N อันดับ

โดยภาพรวมรายละเอียดขั้นตอนวิธีการสร้างรายการแนะนำสถานที่ที่น่าสนใจ NILR ที่ซึ่งประกอบด้วย 1. การคำนวณค่าคะแนนความน่าสนใจของสถานที่ 2. การจัดอันดับรายการแนะนำสถานที่ และ 3. การสร้างรายการแนะนำสถานที่ที่น่าสนใจมากที่สุด N อันดับ แสดงดังภาพที่ 3.6

Algorithm 5 N-most Interesting location-based recommender system

Input: A set of locations, $L = \{l_1, l_2, \ldots, l_q\}$,
A set of users, $U = \{u_1, u_2, \ldots, u_p\}$,
A number of iterations for processing, t
Current location of a new-user nu, $loc(nu)$,
A maximum distance (in km) between current location of the user and a recommended shop, d, and
A number of shops to be recommended, N
Output: An order list of N-most interesting locations, $L^{nu} = \{l_y, \ldots, l_z\}$

ISF^L = HITS(L, U, x)
ISP^L = ReturnHITS(L, U, x)
IL = Ranking(L, ISF^L, ISP^L)
L^{nu} = TopNGeneration($loc(nu), d, N, IL, L$)

ภาพที่ 3.6 ภาพรวมรายละเอียดขั้นตอนวิธีการสร้างรายการแนะนำสถานที่ NILR

3.3 ตัวอย่างการสร้างรายการแนะนำสถานที่ที่น่าสนใจด้วยวิธี NILR

รหัสผู้ใช้	รหัสสถานที่	รหัสหมวดหมู่สถานที่	หมวดหมู่สถานที่	ละติจูด	ลองจิจูด	เวลาในการเช็คอิน	เวลา UTC
u_1	l_3	4bf58dd8d48988d16e941735	Fast Food Restaurant	35.577	19.660	540	Tue Apr 03 21:33:17 +0000 2012
u_2	l_1	4bf58dd8d48988d1c4941735	Restaurant	35.627	39.714	40	Tue Apr 03 21:50:15 +0000 2012
u_3	l_1	4bf58dd8d48988d1e0931735	Coffee Shop	35.689	39.700	40	Tue Apr 03 22:13:21 +0000 2012
u_1	l_4	4bf58dd8d48988d111941735	Japanese Restaurant	35.652	39.544	540	Tue Apr 03 22:23:00 +0000 2012
u_6	l_4	4bf58dd8d48988d1e0931735	Coffee Shop	35.691	39.703	40	Tue Apr 03 22:23:18 +0000 2012
u_4	l_3	4bf58dd8d48988d16e941735	Fast Food Restaurant	35.646	39.745	540	Tue Apr 03 22:25:30 +0000 2012
u_5	l_3	4bf58dd8d48988d1c4941735	Restaurant	35.627	39.7143	40	Tue Apr 03 22:26:33 +0000 2012
u_3	l_3	4d4ae6fc7a7b7dea34424761	Fried Chicken Joint	35.789	39.661	540	Tue Apr 03 22:27:30 +0000 2012
u_2	l_2	4bf58dd8d48988d16e941735	Fast Food Restaurant	35.574	139.658	540	Tue Apr 03 22:27:54 +0000 2012

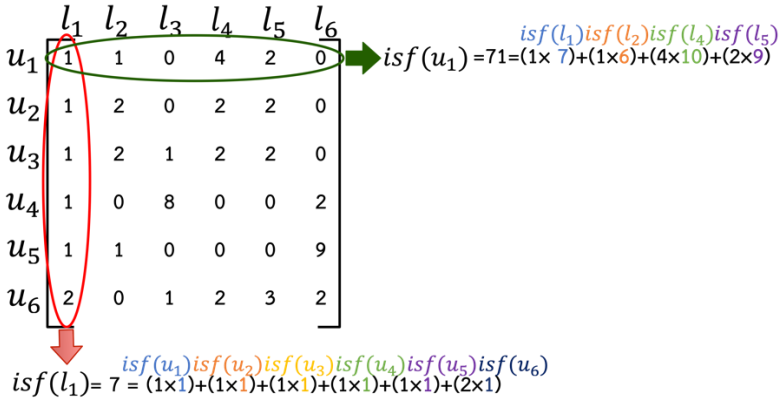
ภาพที่ 3.7 ตัวอย่างข้อมูลประวัติการเช็คอินของผู้ใช้

จากภาพที่ 3.7 แสดงตัวอย่างข้อมูลประวัติการเช็คอินของผู้ใช้ที่ซึ่งประกอบด้วย 1. รหัสผู้ใช้ 2. รหัสสถานที่ 3. รหัสหมวดหมู่สถานที่ 4. หมวดหมู่สถานที่ 5. ละติจูด 6. ลองจิจูด 7. เวลาในการเช็คอิน และ 8. เวลา UTC โดยในขั้นตอนการคำนวณค่าคะแนนความน่าสนใจของสถานที่จำเป็นต้องใช้ 1. รหัสผู้ใช้ 2. รหัสสถานที่และ 3. รหัสหมวดหมู่สถานที่ ที่ซึ่งรหัสผู้ใช้แสดงถึงผู้ใช้ที่ทำการเช็คอินในสถานที่และรหัสสถานที่นั้นแสดงถึงชื่อของสถานที่เมื่อทำการสร้างรายการแนะนำ และในงานนี้ได้มีการพิจารณาการสร้างรายการแนะนำสถานที่ที่น่าสนใจในหมวดหมู่เดียวกัน จากประวัติการเช็คอินของผู้ใช้ในพื้นที่ที่สนใจและหมวดหมู่เดียวกันจะถูกแปลงให้อยู่ในรูปแบบของเมทริกซ์ แสดงดังภาพที่ 3.8 ที่ซึ่งแต่ละแถวแสดงถึงผู้ใช้และแต่ละคอลัมภ์แสดงถึงสถานที่ โดยข้อมูลภายในเมทริกซ์แสดงถึงความถี่ในการเช็คอินของผู้ใช้ เช่นผู้ใช้ u_1 มีการเช็คอินโดยการไปเยี่ยมชมที่สถานที่ l_1, l_2, l_4, l_5 จำนวน 1 ครั้ง 1 ครั้ง 4 ครั้งและ 2 ครั้ง ตามลำดับ

	l_1	l_2	l_3	l_4	l_5	l_6
u_1	1	1	0	4	2	0
u_2	1	2	0	2	2	0
u_3	1	2	1	2	2	0
u_4	1	0	8	0	0	2
u_5	1	1	0	0	0	9
u_6	2	0	1	2	3	2

ภาพที่ 3.8 เมทริกซ์สำหรับการคำนวณค่าคะแนนความน่าสนใจของสถานที่

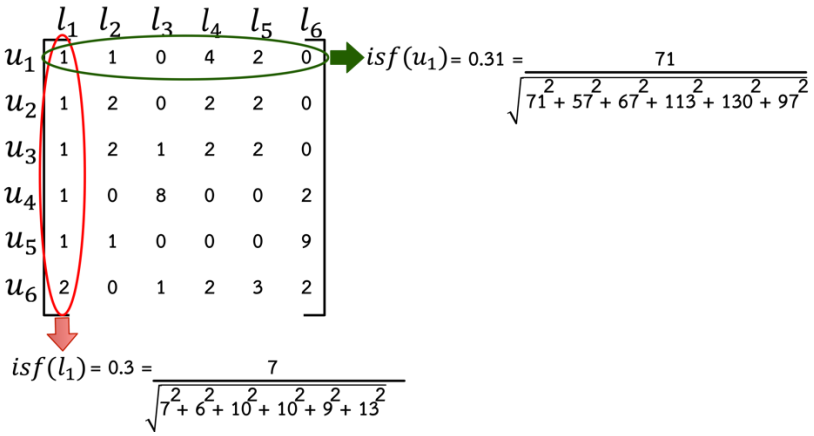
ขั้นตอนแรกจะทำการคำนวณค่าคะแนนความน่าสนใจของสถานที่โดยพิจารณาความถี่ในการเช็คอินของผู้ใช้ ดังนิยามที่ 1 ตัวอย่างเช่น การคำนวณค่าคะแนนความน่าสนใจของสถานที่ l_1 ($isf(l_1)$) จะได้ว่า $isf(l_1)$ มีค่าเท่ากับ 7 คำนวณจาก $isf(l_1) = (1 \times 1) + (1 \times 1) + (1 \times 1) + (1 \times 1) + (1 \times 1) + (2 \times 1)$ เมื่อคำนวณจนครบทุกสถานที่แล้ว จะได้ค่าคะแนนความน่าสนใจของสถานที่ ดังนี้ $isf(l_1)$ มีค่าเท่ากับ 7 $isf(l_2)$ มีค่าเท่ากับ 6 $isf(l_3)$ มีค่าเท่ากับ 10 $isf(l_4)$ มีค่าเท่ากับ 10 $isf(l_5)$ มีค่าเท่ากับ 9 และ $isf(l_6)$ มีค่าเท่ากับ 13 ตามลำดับ จากนั้นในขั้นตอนที่สองจะทำการคำนวณค่าคะแนนของผู้ใช้ ตัวอย่างเช่น การคำนวณค่าคะแนนของผู้ใช้ u_1 มีค่าเท่ากับ 71 โดยคำนวณจาก $isf(u_1) = (1 \times 7) + (1 \times 6) + (4 \times 10) + (2 \times 9)$ เมื่อคำนวณจนครบทุกสถานที่แล้ว จะได้ค่าคะแนนความน่าสนใจของผู้ใช้ ดังนี้ $isf(u_1)$ มีค่าเท่ากับ 71 $isf(u_2)$ มีค่าเท่ากับ 57 $isf(u_3)$ มีค่าเท่ากับ 67 $isf(u_4)$ มีค่าเท่ากับ 113 $isf(u_5)$ มีค่าเท่ากับ 130 และ $isf(u_6)$ มีค่าเท่ากับ 97 ตามลำดับ แสดงตัวอย่างการคำนวณค่าคะแนนความน่าสนใจของสถานที่ l_1 และค่าคะแนนผู้ใช้โดยพิจารณาความถี่ในการเช็คอินของผู้ใช้ u_1 ดังภาพที่ 3.9



ภาพที่ 3.9 ตัวอย่างการคำนวณค่าคะแนนความน่าสนใจของสถานที่ l_1 และค่าคะแนนผู้ใช้ u_1 โดยพิจารณาความถี่ในการเช็คอินของผู้ใช้

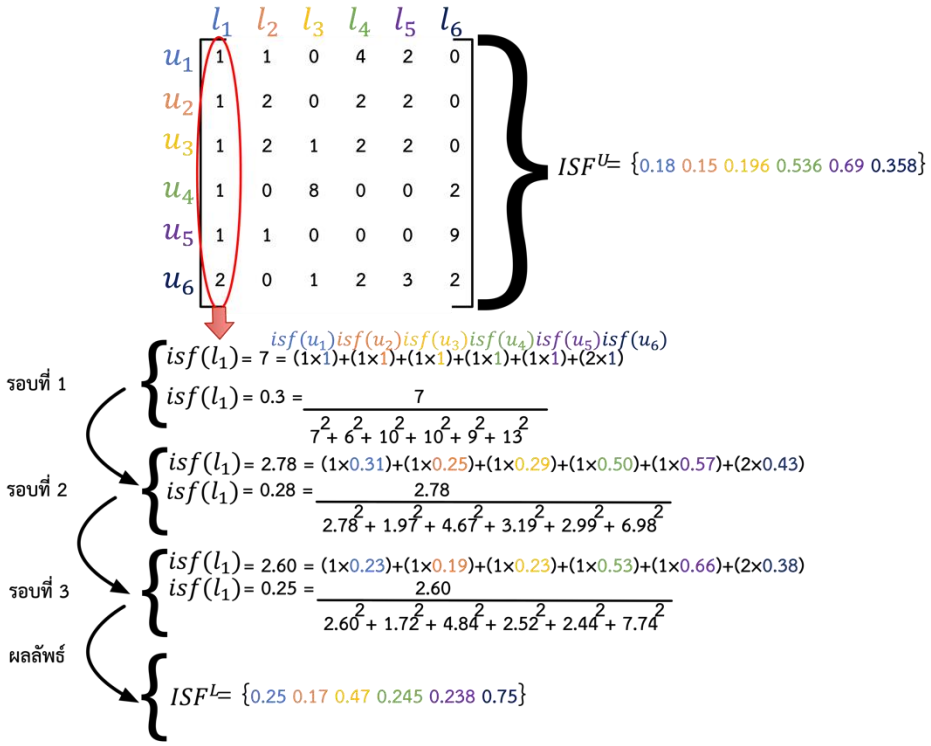
จากนั้นค่าคะแนนความน่าสนใจของสถานที่และค่าคะแนนผู้ใช้จะถูกปรับให้มีจำนวนที่ลดลงโดยใช้ขั้นตอนวิธี *L2*-normalization ตัวอย่างเช่น สถานที่ l_1 ซึ่งมีค่า $isf(l_1)$ มีค่าเท่ากับ 7 จะถูกปรับเป็น 0.30 นั่นคือ $isf(l_1) = \frac{7}{\sqrt{7^2+6^2+10^2+10^2+9^2+13^2}}$ เมื่อคำนวณจนครบทุกสถานที่แล้ว จะได้ค่าคะแนนความน่าสนใจใหม่ของสถานที่ดังนี้ $isf(l_1)$ มีค่าเท่ากับ 0.30 $isf(l_2)$ มีค่าเท่ากับ 0.26 $isf(l_3)$ มีค่าเท่ากับ 0.43 $isf(l_4)$ มีค่าเท่ากับ 0.43 $isf(l_5)$ มีค่า

เท่ากับ 0.39 และ $isf(l_6)$ มีค่าเท่ากับ 0.56 ตามลำดับ ตัวอย่างการคำนวณค่าคะแนนผู้ใช้สามารถคำนวณได้โดยผู้ใช้ u_1 มีค่า $isf(u_1)$ เท่ากับ 0.31 สามารถคำนวณ นั่นคือ $isf(u_1) = \frac{71}{\sqrt{71^2+57^2+67^2+113^2+130^2+97^2}}$ และเมื่อคำนวณจนครบทุกผู้ใช้แล้ว จะได้ค่าคะแนนความน่าสนใจใหม่ของผู้ใช้ดังนี้ $isf(u_1)$ มีค่าเท่ากับ 0.31 $isf(u_2)$ มีค่าเท่ากับ 0.25 $isf(u_3)$ มีค่าเท่ากับ 0.29 $isf(u_4)$ มีค่าเท่ากับ 0.50 $isf(u_5)$ มีค่าเท่ากับ 0.57 และ $isf(u_6)$ มีค่าเท่ากับ 0.43 ตามลำดับ แสดงตัวอย่างการปรับลดค่าคะแนน $isf(l_1)$ และ $isf(u_1)$ โดยใช้ *L2*-normalization ดังภาพที่ 3.10



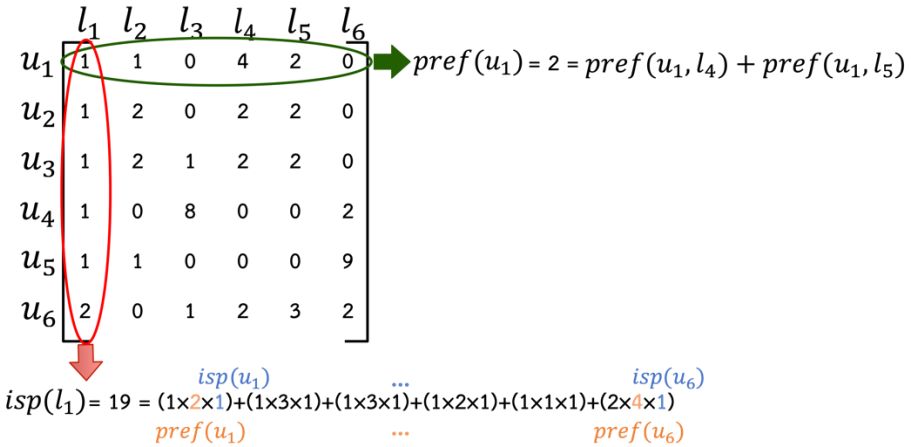
ภาพที่ 3.10 ตัวอย่างการปรับลดค่าคะแนน $isf(l_1)$ และ $isf(u_1)$ โดยใช้ *L2*-normalization

เมื่อทำการคำนวณแบบวนซ้ำเพื่ออัพเดทค่าคะแนนความน่าสนใจของสถานที่และค่าคะแนนของผู้ใช้ จะได้ค่าคะแนนสุดท้ายเพื่อเตรียมพร้อมสำหรับการแนะนำ ดังภาพที่ 3.11 แสดงตัวอย่างการคำนวณค่าคะแนนความน่าสนใจของสถานที่ l_1 และค่าคะแนนของผู้ใช้เมื่ออัพเดทค่าคะแนนทั้งหมด 3 รอบ



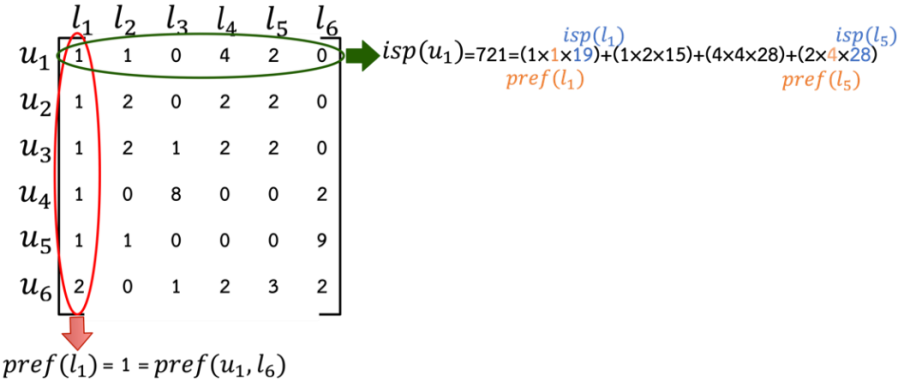
ภาพที่ 3.11 ตัวอย่างการคำนวณค่าคะแนนความน่าสนใจของสถานที่ l_1 และค่าคะแนนของผู้ใช้เมื่ออัพเดทค่าคะแนนทั้งหมด 3 รอบ

ต่อมาค่าคะแนนความน่าสนใจของสถานที่โดยพิจารณาค่าความชอบของผู้ใช้และค่าคะแนนของผู้ใช้โดยพิจารณาค่าความชอบของสถานที่จะถูกคำนวณ ดังนิยามที่ 2 ในขั้นตอนแรกค่าความชอบของผู้ใช้จะถูกพิจารณา ตัวอย่างเช่น ค่าความชอบของผู้ใช้ u_1 ($pref(u_1)$) มีค่าเท่ากับ 2 โดยคำนวณจาก $pref(u_1) = 2$ ($pref(u_1, l_4) + pref(u_1, l_5)$) เมื่อคำนวณจนครบทุกผู้ใช้แล้ว จะได้ค่าความชอบของผู้ใช้ดังนี้ $pref(u_1)$ มีค่าเท่ากับ 2 $pref(u_2)$ มีค่าเท่ากับ 3 $pref(u_3)$ มีค่าเท่ากับ 3 $pref(u_4)$ มีค่าเท่ากับ 2 $pref(u_5)$ มีค่าเท่ากับ 1 และ $pref(u_6)$ มีค่าเท่ากับ 4 ตามลำดับ จากนั้นค่าคะแนนความน่าสนใจของสถานที่โดยพิจารณาค่าความชอบของผู้ใช้ ตัวอย่างเช่น ค่าคะแนนความน่าสนใจของสถานที่ l_1 ($isp(l_1)$) มีค่าเท่ากับ 19 นั่นคือ $isp(l_1) = (1 \times 2 \times 1) + (1 \times 3 \times 1) + (1 \times 3 \times 1) + (1 \times 2 \times 1) + (1 \times 1 \times 1) + (2 \times 4 \times 1)$ เมื่อคำนวณจนครบทุกสถานที่แล้ว จะได้ค่าคะแนนดังนี้ $isp(l_1)$ มีค่าเท่ากับ 19 $isp(l_2)$ มีค่าเท่ากับ 15 $isp(l_3)$ มีค่าเท่ากับ 23 $isp(l_4)$ มีค่าเท่ากับ 28 $isp(l_5)$ มีค่าเท่ากับ 28 และ $isp(l_6)$ มีค่าเท่ากับ 21 ตามลำดับ แสดงตัวอย่างการคำนวณค่าความชอบของผู้ใช้ u_1 และค่าคะแนนความน่าสนใจของสถานที่ l_1 ดังภาพที่ 3.12



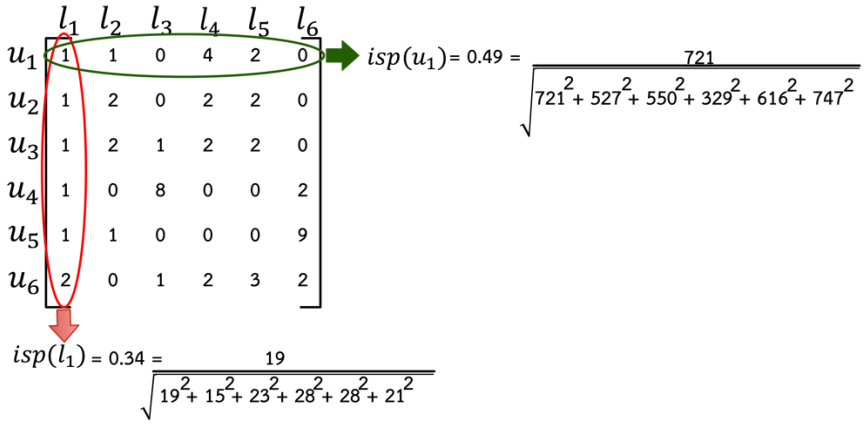
ภาพที่ 3.12 ตัวอย่างการคำนวณค่าความชอบของผู้ใช้ u_1 และค่าคะแนนความน่าสนใจของสถานที่ l_1

ต่อมาค่าความชอบของสถานที่จะถูกพิจารณา ตัวอย่างเช่น ค่าความชอบของผู้ใช้ l_1 ($pref(l_1)$) มีค่าเท่ากับ 1 โดยคำนวณจาก $pref(l_1) = 1$ ($pref(u_6, l_1)$) เมื่อคำนวณจนครบทุกสถานที่แล้ว จะได้ค่าความชอบของสถานที่ดังนี้ $pref(l_1)$ มีค่าเท่ากับ 1 $pref(l_2)$ มีค่าเท่ากับ 2 $pref(l_3)$ มีค่าเท่ากับ 1 $pref(l_4)$ มีค่าเท่ากับ 4 $pref(l_5)$ มีค่าเท่ากับ 4 และ $pref(l_6)$ มีค่าเท่ากับ 3 ตามลำดับ จากนั้นค่าความชอบของผู้ใช้โดยพิจารณาค่าคะแนนความน่าสนใจของสถานที่ ตัวอย่างเช่น ค่าคะแนนของผู้ใช้ u_1 ($isp(u_1)$) มีค่าเท่ากับ 721 โดยคำนวณจาก $isp(u_1) = (1\times1\times19)+(1\times2\times15)+(4\times4\times28)+(2\times4\times28)$ เมื่อคำนวณจนครบทุกผู้ใช้แล้ว จะได้ค่าคะแนนของผู้ใช้ดังนี้ $isp(u_1)$ มีค่าเท่ากับ 721 $isp(u_2)$ มีค่าเท่ากับ 527 $isp(u_3)$ มีค่าเท่ากับ 550 $isp(u_4)$ มีค่าเท่ากับ 329 $isp(u_5)$ มีค่าเท่ากับ 616 และ $isp(u_6)$ มีค่าเท่ากับ 747 ตามลำดับ แสดงดังภาพที่ 3.13



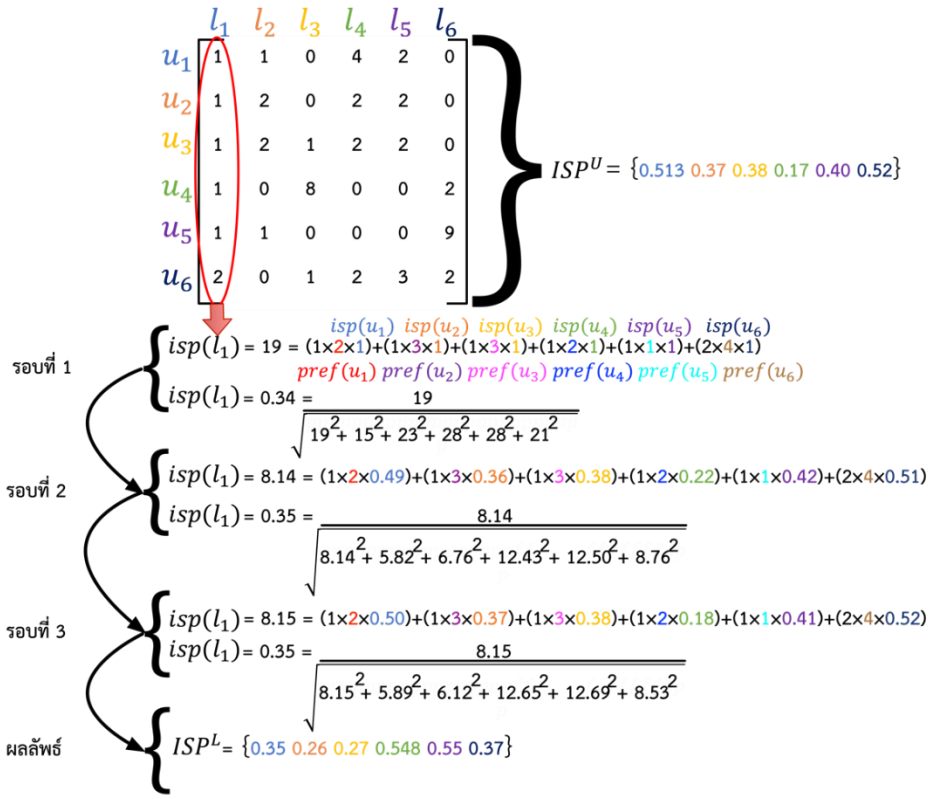
ภาพที่ 3.13 ตัวอย่างการคำนวณค่าความชอบของสถานที่ l_1 และค่าคะแนนของผู้ใช้ u_1

จากนั้นขั้นตอนวิธี *L2*-normalization จะถูกใช้ในการปรับค่าคะแนนคะแนนความน่าสนใจของสถานที่และค่าคะแนนของผู้ใช้ ตัวอย่างเช่น สถานที่ l_1 ซึ่งมีค่า $isp(l_1)$ มีค่าเท่ากับ 19 จะถูกปรับเป็น 0.34 นั่นคือ $isp(l_1) = \frac{19}{\sqrt{19^2+15^2+23^2+28^2+28^2+21^2}}$ เมื่อคำนวณจนครบทุกสถานที่แล้ว จะได้ค่าคะแนนความน่าสนใจใหม่ของสถานที่ ดังนี้ $isp(l_1)$ มีค่าเท่ากับ 0.34 $isp(l_2)$ มีค่าเท่ากับ 0.27 $isp(l_3)$ มีค่าเท่ากับ 0.41 $isp(l_4)$ มีค่าเท่ากับ 0.50 $isp(l_5)$ มีค่าเท่ากับ 0.50 และ $isp(l_6)$ มีค่าเท่ากับ 0.38 ตามลำดับ สำหรับตัวอย่างการคำนวณค่าคะแนนผู้ใช้สามารถคำนวณได้โดยผู้ใช้ u_1 มีค่า $isp(u_1)$ เท่ากับ 721 จะถูกปรับเป็น 0.49 คำนวณจาก $isp(u_1) = \frac{721}{\sqrt{721^2+527^2+550^2+329^2+616^2+747^2}}$ และเมื่อคำนวณจนครบทุกผู้ใช้แล้ว จะได้ค่าคะแนนความชอบของผู้ใช้ ดังนี้ $isp(u_1)$ มีค่าเท่ากับ 0.49 $isp(u_2)$ มีค่าเท่ากับ 0.36 $isp(u_3)$ มีค่าเท่ากับ 0.38 $isp(u_4)$ มีค่าเท่ากับ 0.22 $isp(u_5)$ มีค่าเท่ากับ 0.42 และ $isp(u_6)$ มีค่าเท่ากับ 0.51 ตามลำดับ แสดงดังภาพที่ 3.14



ภาพที่ 3.14 ตัวอย่างการปรับลดค่าคะแนน $isp(l_1)$ และ $isp(u_1)$โดยใช้ *L2*-normalization

เมื่อทำการคำนวณแบบวนซ้ำเพื่ออัพเดทค่าคะแนนความน่าสนใจของสถานที่และค่าคะแนนของผู้ใช้โดยพิจารณาค่าความชอบ จะได้ค่าคะแนนสุดท้ายเพื่อเตรียมพร้อมสำหรับการแนะนำ ดังภาพที่ 3.15 แสดงค่าคะแนนความน่าสนใจของสถานที่และค่าคะแนนของผู้ใช้เมื่ออัพเดทค่าคะแนนทั้งหมด 3 รอบ



ภาพที่ 3.15 ตัวอย่างค่าคะแนนของสถานที่ l_1 และค่าคะแนนของผู้ใช้โดยพิจารณาค่าความชอบเมื่อคำนวณจนครบ 3 รอบ

สำหรับวิธีการในการจัดอันดับสถานที่ที่น่าสนใจ ในขั้นตอนแรกจะทำการจัดอันดับค่าคะแนนสถานที่ของทั้งสองรายการจากค่ามากไปหาค่าน้อย ซึ่งจะได้รายการสถานที่ที่น่าสนใจเมื่อพิจารณาการเช็คอินของผู้ใช้ (ISF^L) คือ $l_6\ l_3\ l_1\ l_4\ l_5$ และ l_2 โดยพิจารณาจากค่าคะแนน 0.75 0.47 0.25 0.245 0.24 และ 0.17 ตามลำดับ สำหรับรายการสถานที่ที่น่าสนใจเมื่อพิจารณาค่าความชอบของผู้ใช้ (ISP^L) จะได้รายการสถานที่ที่เรียงลำดับดังนี้ $l_5\ l_4\ l_6\ l_1\ l_3$ และ l_2 โดยพิจารณาจากค่าคะแนน 0.55 0.548 0.37 0.35 0.27 และ 0.26 ตามลำดับ จากนั้นอันดับที่มากที่สุดทั้งสองรายการโดยพิจารณาจากค่าคะแนน max^{freq} และ max^{pref} จะถูกพิจารณานั่นคือ สถานที่ l_6 และสถานที่ l_5 ที่ซึ่งค่าคะแนนที่มากที่สุดคือ l_6 ดังนั้นเราจะได้รายการแนะนำสถานที่ที่น่าสนใจสุดท้าย (IL) สำหรับอันดับ 1 และ 2 คือ $IL = \{l_6 \succ l_5\}$ จากนั้นสถานที่ l_6 และ l_5 จะถูกลบออกจากรายการที่พิจารณา ISF^L และ ISP^L โดยกระบวนการนี้จะถูกคำนวณซ้ำจนกระทั่งทุกสถานที่ได้รับการพิจารณา ดังนั้นรายการแนะนำสถานที่ที่น่าสนใจสุดท้าย IL คือ $\{l_6 \succ l_5 \succ l_4 \succ l_3 \succ l_1 \succ l_2\}$ แสดงตัวอย่างวิธีการจัดอันดับรายการแนะนำสถานที่ที่น่าสนใจ ดังภาพที่ 3.16

$$ISF^L = \{\underset{0.25}{l_1}\ \underset{0.17}{l_2}\ \underset{0.47}{l_3}\ \underset{0.245}{l_4}\ \underset{0.24}{l_5}\ \underset{0.75}{l_6}\}$$

$$ISP^L = \{0.35\ 0.26\ 0.27\ 0.548\ 0.55\ 0.37\}$$

จัดอันดับสถานที่ที่น่าสนใจ

$$ISF^L = \{\underset{0.75}{l_6}\ \underset{0.47}{l_3}\ \underset{0.25}{l_1}\ \underset{0.245}{l_4}\ \underset{0.24}{l_5}\ \underset{0.17}{l_2}\}$$

$$ISP^L = \{\underset{0.55}{l_5}\ \underset{0.548}{l_4}\ \underset{0.37}{l_6}\ \underset{0.35}{l_1}\ \underset{0.27}{l_3}\ \underset{0.26}{l_2}\}$$

สร้างรายการแนะนำสถานที่ที่น่าสนใจสุดท้าย

$$\{max^{freq} > max^{pref}\}$$

$$IL = \{l_6 \succ l_5\}$$

ลบสถานที่ที่พิจารณาแล้ว

$$ISF^L = \{\underset{0.47}{l_3}\ \underset{0.25}{l_1}\ \underset{0.245}{l_4}\ \underset{0.24}{l_5}\ \underset{0.17}{l_2}\}$$

$$ISP^L = \{\underset{0.548}{l_4}\ \underset{0.37}{l_6}\ \underset{0.35}{l_1}\ \underset{0.27}{l_3}\ \underset{0.26}{l_2}\}$$

$$\vdots$$

$$IL = \{l_6 \succ l_5 \succ l_4 \succ l_3 \succ l_1 \succ l_2\}$$

ภาพที่ 3.16 ตัวอย่างวิธีการจัดอันดับรายการแนะนำสถานที่ที่น่าสนใจ

สำหรับการแนะนำสถานที่ที่น่าสนใจให้กับผู้ใช้ ตัวอย่างเช่น เมื่อผู้ใช้ใหม่ nu ต้องการรายการแนะนำสถานที่ที่น่าสนใจที่ซึ่งสถานที่ $l_6\ l_4\ l_3\ l_1$ และ l_2 ตั้งอยู่ในระแวกใกล้เคียงกับที่อยู่ของผู้ใช้และผู้ใช้ nu ต้องการรายการสถานที่ที่น่าสนใจ 3 อันดับที่มากที่สุดจะได้ว่ารายการแนะนำสถานที่ที่น่าสนใจ L^{nu} คือ $\{l_6 \succ l_4 \succ l_3\}$ แสดงตัวอย่างวิธีการสร้างรายการแนะนำสถานที่ที่น่าสนใจให้กับผู้ใช้ nu 3 อันดับที่น่าสนใจที่สุด ดังภาพที่ 3.17

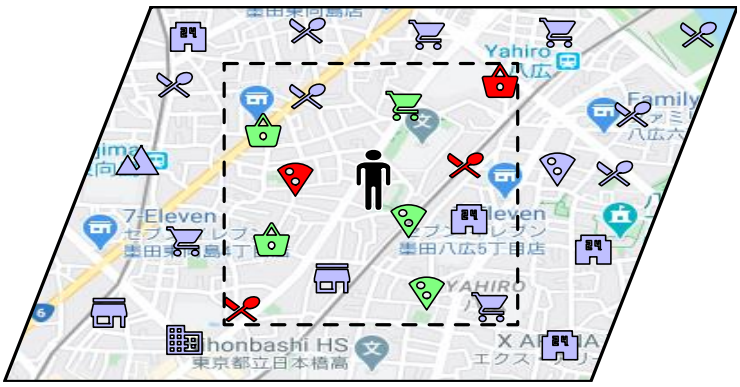
ภาพที่ 3.17 ตัวอย่างวิธีการสร้างรายการแนะนำสถานที่ที่น่าสนใจให้กับผู้ใช้ nu 3 อันดับที่น่าสนใจที่สุด

บทที่ 4
ผลการดำเนินงาน

ในบทนี้จะกล่าวถึง การตั้งค่าการทดลอง รายละเอียดต่าง ๆ ของการเตรียมข้อมูลประวัติการเช็คอินของผู้ใช้และรายละเอียดวิธีการที่ใช้ในการวัดประสิทธิภาพความถูกต้อง นั่นคือ ขั้นตอนวิธี Precision Recall และ Average ranking และผลการทดลองของขั้นตอนวิธีที่นำเสนอโดยเปรียบเทียบกับขั้นตอนวิธี HITS

4.1 การตั้งค่าการทดลองและวิธีการที่ใช้ในการวัดประสิทธิภาพความถูกต้อง

เนื่องจากฐานข้อมูลมีขนาดใหญ่และมีความเบาบางสูง ดังนั้นในงานวิจัยนี้จึงได้มีการเตรียมข้อมูลโดยพิจารณาผู้ใช้ซึ่งมีการเช็คอินมากกว่า 1 สถานที่ และสถานที่ต้องถูกเช็คอินโดยผู้ใช้มากกว่า 1 ผู้ใช้ โดยชุดข้อมูลในการทดสอบจะถูกแบ่งชุดข้อมูลฝึกฝนและเลือกใช้วิธีการสุ่มข้อมูลแบบเที่ยงตรง 5 กลุ่ม (5-fold cross validation) โดยผู้ใช้ทดสอบที่อยู่ในข้อมูลชุดทดสอบจะถูกพิจารณาเป็นผู้ใช้ใหม่ โดยวิธีการที่นำเสนอนี้จะทำการเปรียบเทียบกับขั้นตอนวิธี HITS (Bao et al., 2012)(Bagci & Karagoz, 2016) เพราะว่าวิธีการนี้มีแนวคิดเหมือนกับวิธีการที่เรานำเสนอ ในผลการทดลองที่นำเสนอนี้จะทำการทดลองโดยพิจารณาจำนวนรายการแนะนำตั้งตั้ง 1 รายการ (N=1) จนถึง 10 รายการ (N=10) ดังนั้นผู้ใช้ที่พิจารณาเป็นข้อมูลทดสอบจะต้องมีสถานที่ในขอบเขตที่พิจารณามากกว่า 10 สถานที่ ที่ซึ่งขอบเขตของสถานที่ที่ถูกพิจารณาเป็นพื้นที่ปัจจุบันของผู้ใช้ได้ใช้วิธี Minimum Bounding Rectangle (MBR) (Bao et al., 2012)(Ying et al., 2017) ดังภาพที่ 4.1 ที่สถานที่สีแดงคือสถานที่ทดสอบซึ่งเป็นสถานที่ที่ใช้ในการกำหนดขอบเขตการพิจารณา (กำหนดให้สถานที่ทดสอบต้องถูกไปโดยผู้ใช้ทดสอบมากกว่า 1 ครั้งถึงแสดงถึงความชอบของผู้ใช้) สถานที่สีเขียวคือสถานที่ที่ถูกแนะนำไปยังผู้ใช้และสถานที่สีฟ้าคือสถานที่ทั่วไป



ภาพที่ 4.1 ตัวอย่างการจำลองสถานที่ปัจจุบันของผู้ใช้โดยใช้วิธี MBR

สำหรับวิธีการที่ใช้ในการวัดประสิทธิภาพความถูกต้องของรายการแนะนำในงานวิจัยนี้ได้ใช้ทั้งหมด 3 วิธีการวัดประกอบด้วยวิธี Precision และวิธี Recall สำหรับการวัดความถูกต้องของรายการแนะนำและวิธี Average ranking สำหรับการวัดความถูกต้องของอันดับของรายการแนะนำ ดังสมการที่ 4.1 4.2 และ 4.3 ตามลำดับ

$$Precision = \frac{\#correct_retrieved}{N} \tag{4.1}$$

$$Recall = \frac{\#correct_retrieved}{\#relevant} \tag{4.2}$$

$$Average\ ranking = \frac{\sum_{i=0}^{n} rank_i}{\#testloc} \tag{4.3}$$

กำหนดให้ $\#correct_retrieved$ คือ จำนวนความถูกต้องของรายการแนะนำ

N คือ จำนวนของรายการแนะนำ

$\#relevant$ คือ จำนวนของสถานที่ทดสอบทั้งหมด

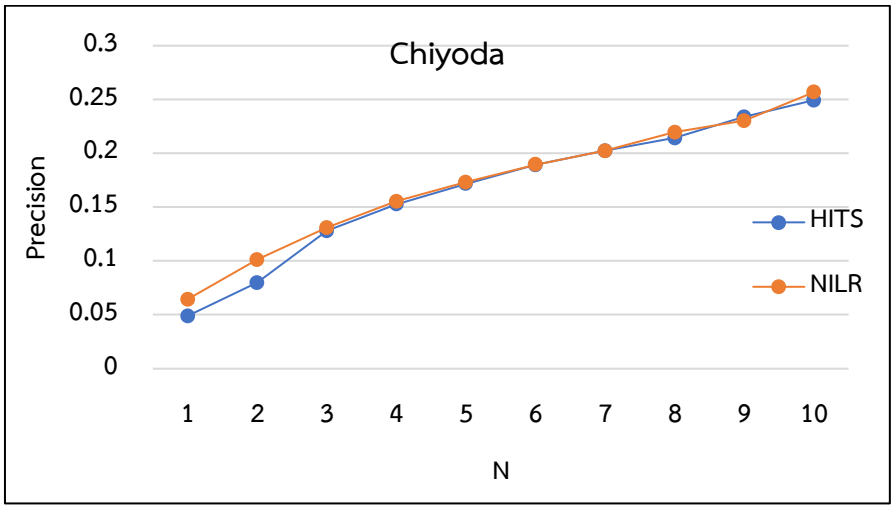
$rank$ คือ อันดับของรายการแนะนำ

$\#testloc$ คือ จำนวนทั้งหมดที่ปรากฏในรายการแนะนำของทั้งสองวิธีการ

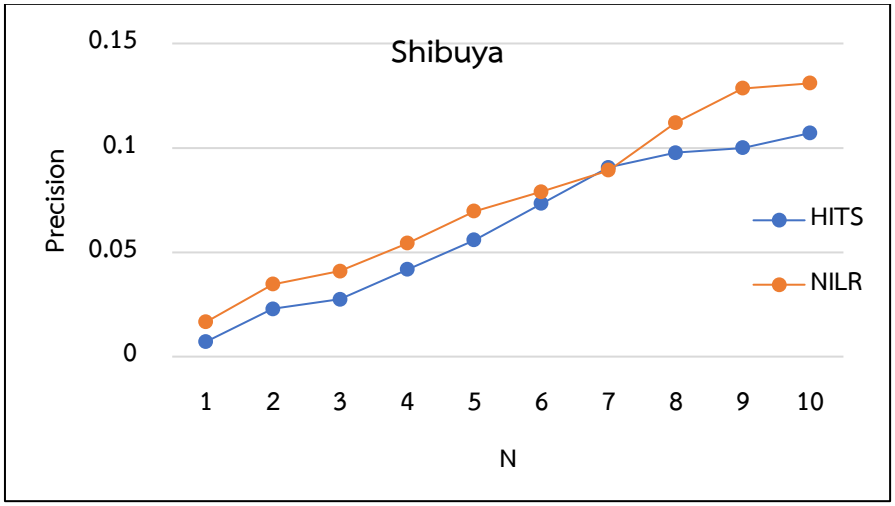
4.2 ผลการทดลอง

ในการทดลองการวัดประสิทธิภาพของวิธีการที่นำเสนอ โดยใช้วิธีการวัดประสิทธิภาพ Precision Recall และ Average ranking ในภาพที่ 4.2 ถึง 4.16 ตามลำดับ

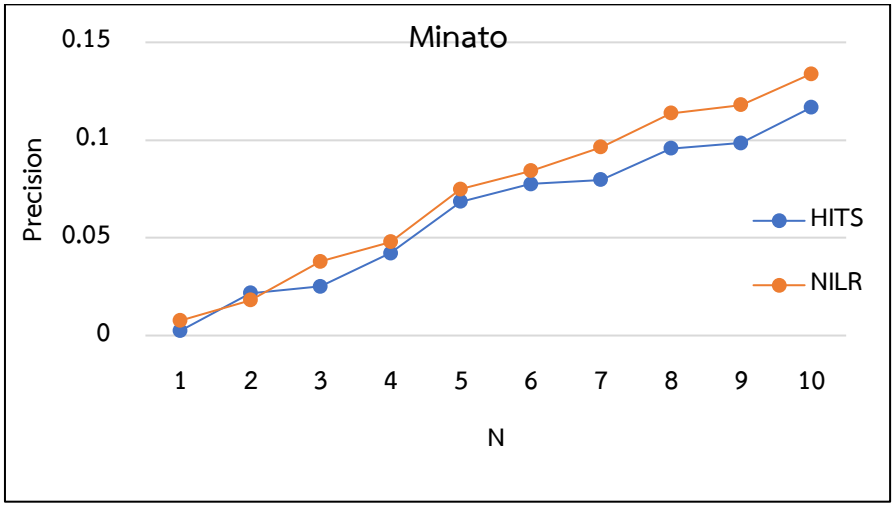
จากภาพที่ 4.2 และภาพที่ 4.11 แสดงให้เห็นว่าวิธีการที่นำเสนอมีความถูกต้องในการสร้างรายการแนะนำสถานที่ที่น่าสนใจสูงกว่า HITS based model ในเมือง Chiyoda Shibuya Minato และ Chuo ยกเว้นเมือง Shinjuku โดยใช้วิธีการวัดประสิทธิภาพ Precision และ Recall ที่ซึ่งวิธีการที่นำเสนอสามารถให้ความถูกต้องที่ดีกว่าเนื่องจากเมืองส่วนมาก ผู้ใช้ในชุดทดสอบมักมีการเยี่ยมชมในสถานที่มีความโดดเด่นทางด้านการกลับมาเที่ยมชมอีกครั้ง แต่ในเมือง Shinjuku วิธี HITS based model มีความถูกต้องในการสร้างรายการแนะนำสถานที่ที่น่าสนใจมากกว่าเนื่องจากผู้ใช้ของเมืองนี้มักไปสถานที่ที่มีความโดดเด่นทางด้านความถี่มากกว่าสถานที่ที่มีความโดดเด่นทางด้านการกลับมาเที่ยมชมอีกครั้ง



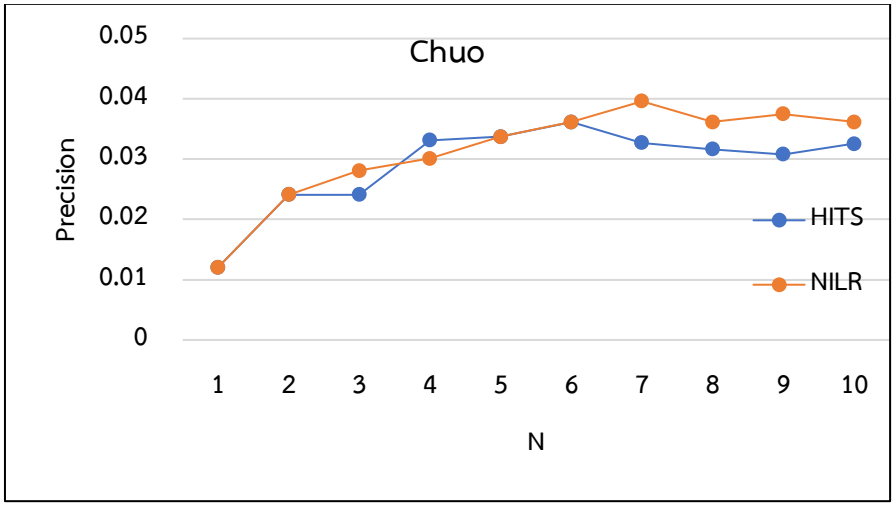
ภาพที่ 4.2 ผลความถูกต้องโดยใช้วิธีการวัดประสิทธิภาพ Precision โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Chiyoda



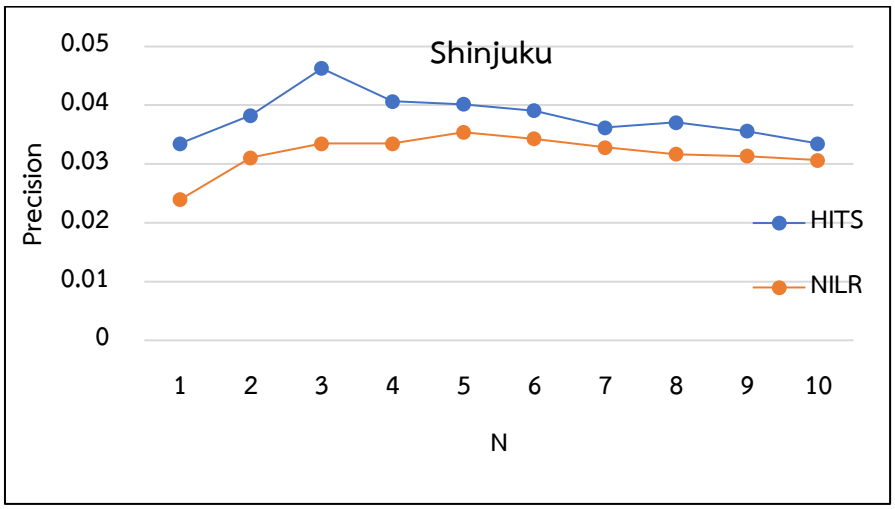
ภาพที่ 4.3 ผลความถูกต้องโดยใช้วิธีการวัดประสิทธิภาพ Precision โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Shibuya



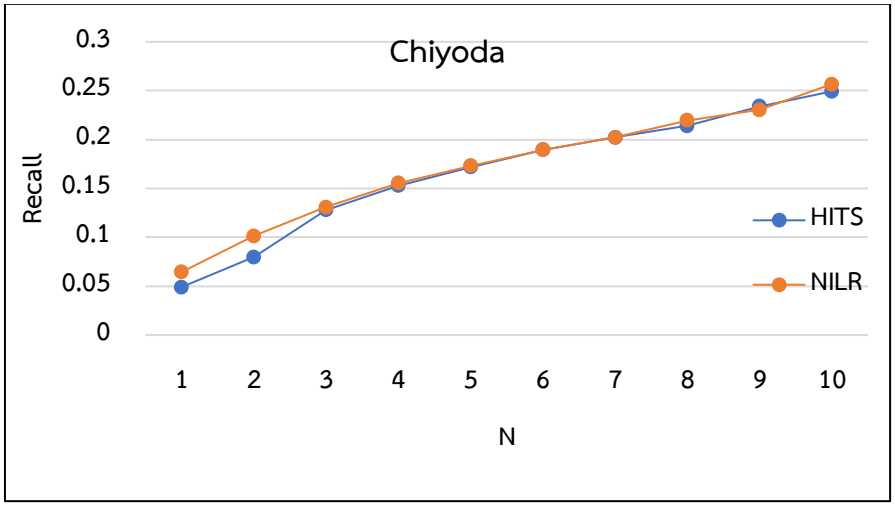
ภาพที่ 4.4 ผลความถูกต้องโดยใช้วิธีการวัดประสิทธิภาพ Precision โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Minato



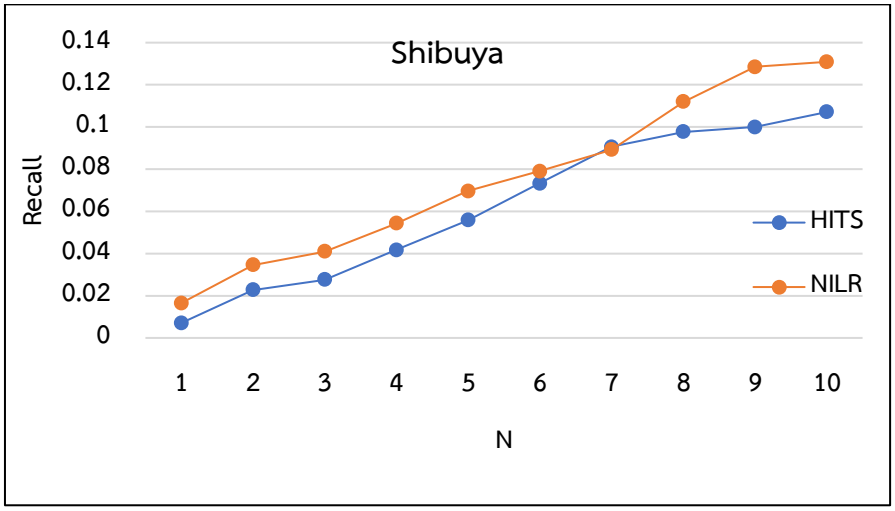
ภาพที่ 4.5 ผลความถูกต้องโดยใช้วิธีการวัดประสิทธิภาพ Precision โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Chuo



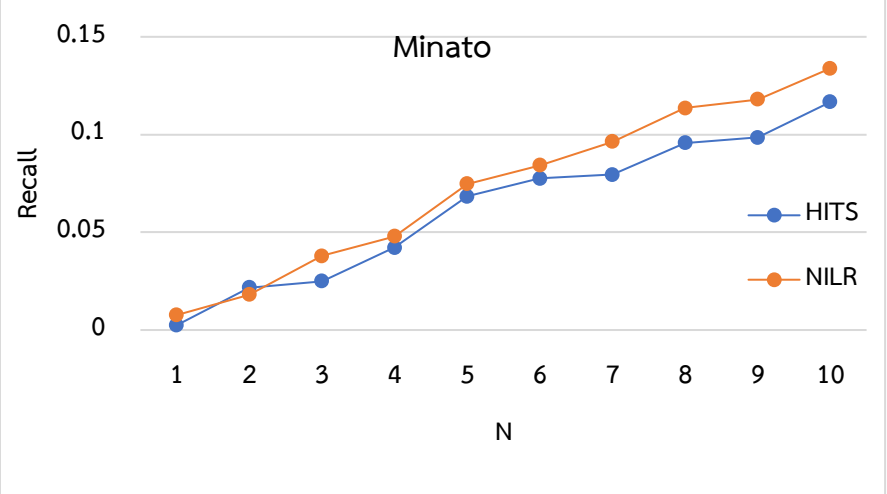
ภาพที่ 4.6 ผลความถูกต้องโดยใช้วิธีการวัดประสิทธิภาพ Precision โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Shinjuku



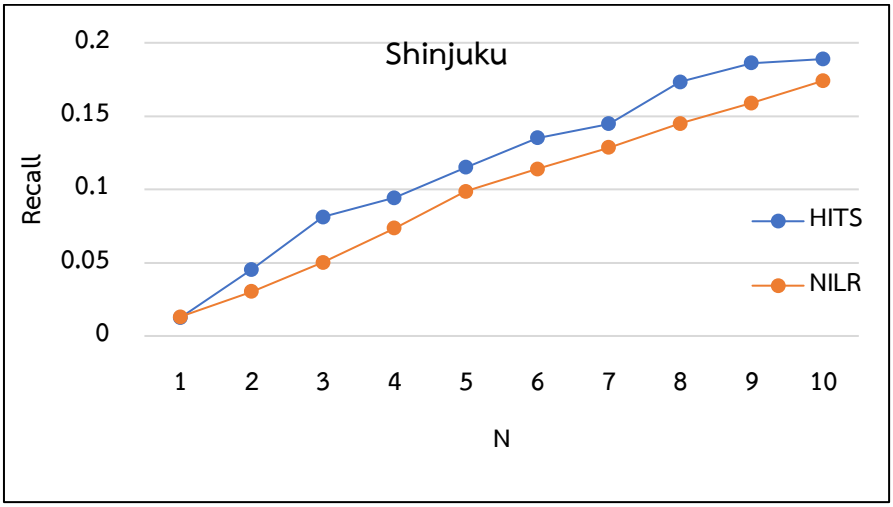
ภาพที่ 4.7 ผลความถูกต้องโดยใช้วิธีการวัดประสิทธิภาพ Recall โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Chiyoda



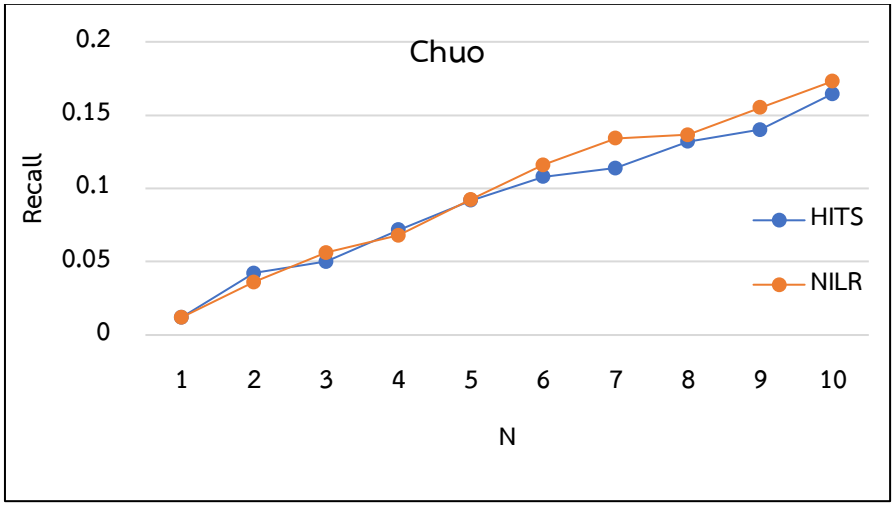
ภาพที่ 4.8 ผลความถูกต้องโดยใช้วิธีการวัดประสิทธิภาพ Recall โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Shibuya



ภาพที่ 4.9 ผลความถูกต้องโดยใช้วิธีการวัดประสิทธิภาพ Recall โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Minato

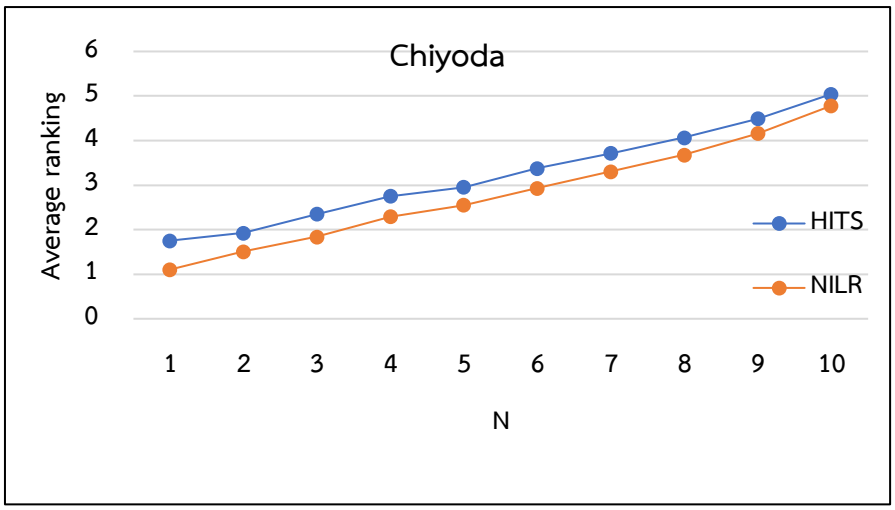


ภาพที่ 4.10 ผลความถูกต้องโดยใช้วิธีการวัดประสิทธิภาพ Recall โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Shinjuku

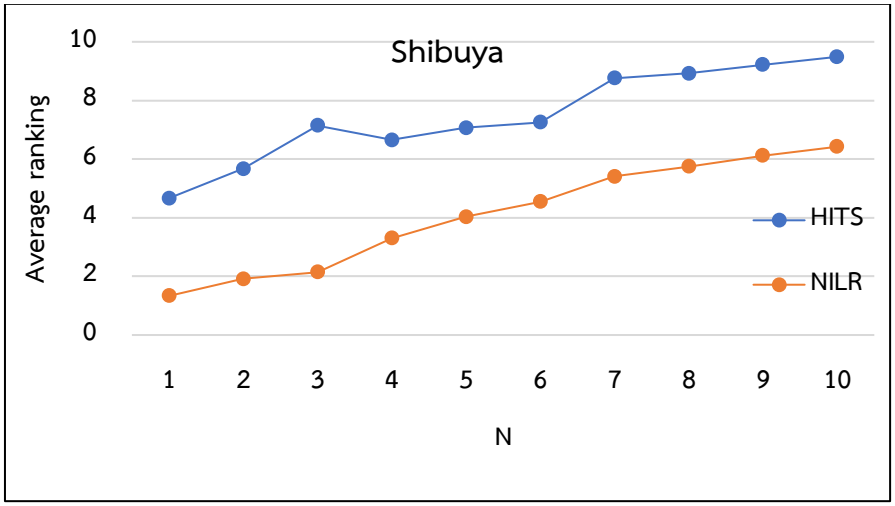


ภาพที่ 4.11 ผลความถูกต้องโดยใช้วิธีการวัดประสิทธิภาพ Recall โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Chuo

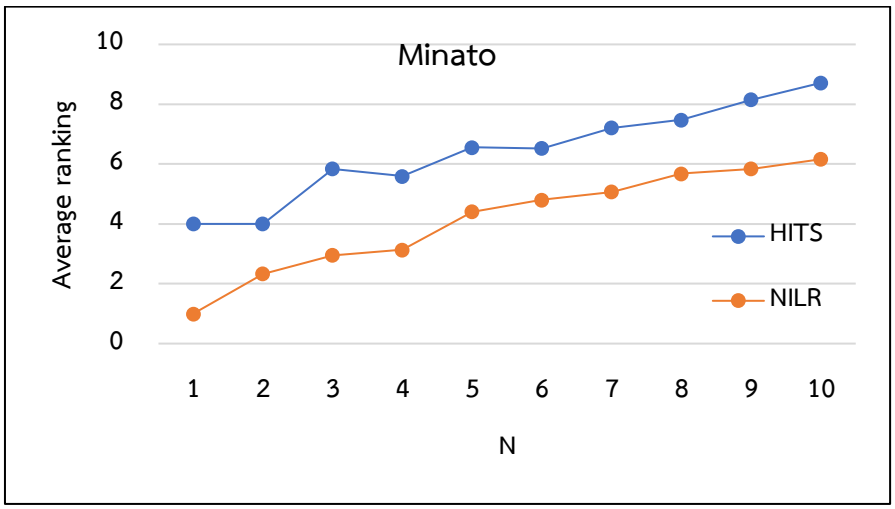
จากภาพที่ 4.12 และภาพที่ 4.16 แสดงให้เห็นว่าวิธีการที่นำเสนอมีความถูกต้องในการจัดอันดับที่ดีกว่าในทุกเมือง สำหรับวิธีการวัดประสิทธิภาพ Average ranking แสดงให้เห็นว่าวิธีการ NILR มีค่าเฉลี่ยของอันดับที่แนะนำได้ถูกต้องอยู่ที่ประมาณช่วง 2-5 อันดับและวิธี HITS based model มีค่าเฉลี่ยของอันดับที่แนะนำได้ถูกต้องอยู่ที่ประมาณช่วง 3-8 อันดับ จากผลการทดลองแสดงให้เห็นว่าวิธีการ NILR สามารถให้ความถูกต้องในการจัดอันดับที่ดีกว่าในทุกเมือง



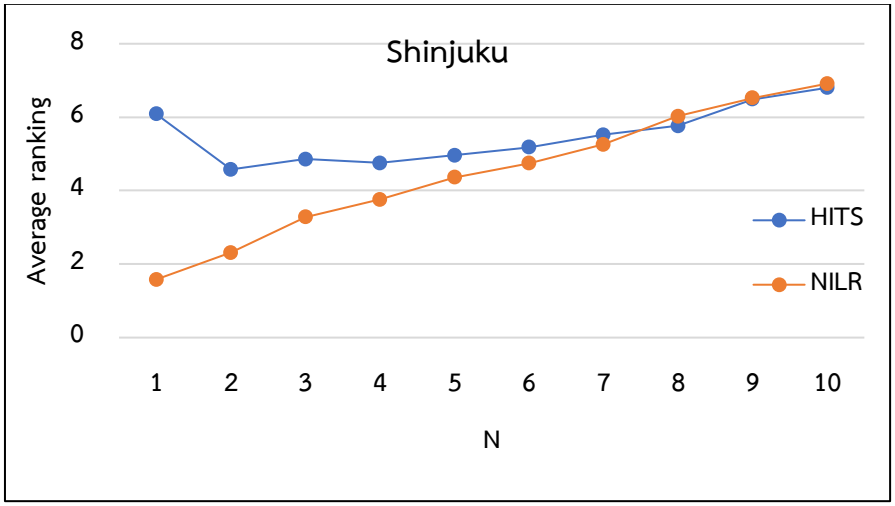
ภาพที่ 4.12 ผลความถูกต้องของการจัดอันดับโดยใช้วิธีการวัดประสิทธิภาพ Average ranking โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Chiyoda



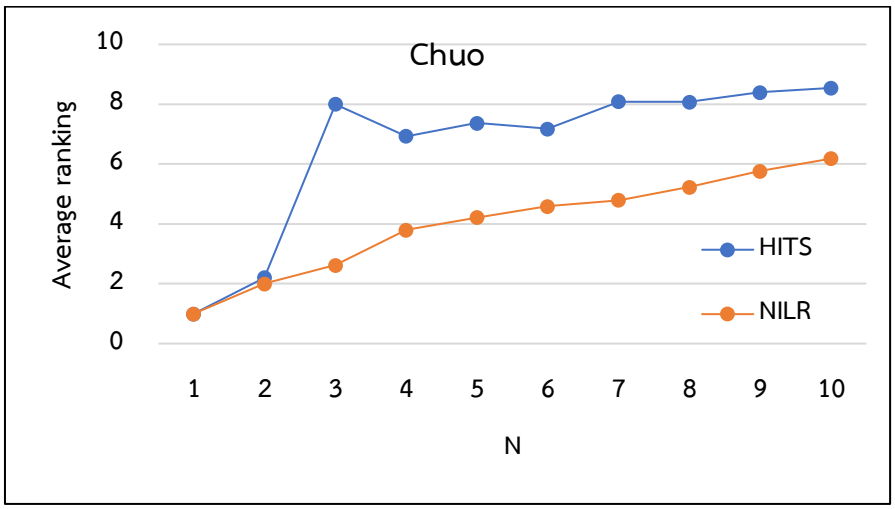
ภาพที่ 4.13 ผลความถูกต้องของการจัดอันดับโดยใช้วิธีการวัดประสิทธิภาพ Average ranking โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Shibuya



ภาพที่ 4.14 ผลความถูกต้องของการจัดอันดับโดยใช้วิธีการวัดประสิทธิภาพ Average ranking โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Minato



ภาพที่ 4.15 ผลความถูกต้องของการจัดอันดับโดยใช้วิธีการวัดประสิทธิภาพ Average ranking โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Shinjuku



ภาพที่ 4.16 ผลความถูกต้องของการจัดอันดับโดยใช้วิธีการวัดประสิทธิภาพ Average ranking โดยพิจารณารายการแนะนำที่ 1 ถึง 10 รายการสถานที่ของเมือง Chuo

สรุปผลในการวัดประสิทธิภาพการทดลองจากค่าเฉลี่ยของรายการแนะนำที่ 1 ถึง 10 รายการสถานที่แสดงให้เห็นว่าวิธีการ NILR มีความถูกต้องที่ดีกว่าวิธี HITS based model ที่ 4.49% and 3.05% โดยใช้วิธี Precision และ Recall ตามลำดับและ NILR มีความถูกต้องในการจัดอันดับที่ดีกว่าวิธี HITS based model ที่ 33.09% โดยใช้วิธี Average ranking ตามลำดับ

บทที่ 5
สรุปและอภิปรายผล

ในงานวิจัยนี้ได้นำเสนอวิธีการสร้างรายการแนะนำสถานที่ที่น่าสนใจเรียกว่า “ระบบแนะนำสถานที่ที่น่าสนใจมากที่สุดเอ็นอันดับ” (N-most Interesting location-based recommender system, NILR) สำหรับแนะนำให้กับผู้ใช้ใหม่ที่ไม่มีความเชี่ยวชาญในพื้นที่ ที่ซึ่งวิธีการที่นำเสนอประกอบไปด้วย 3 ขั้นตอนหลัก คือ 1. การคำนวณค่าคะแนนความน่าสนใจของสถานที่ โดยการปรับปรุงขั้นตอนวิธี HITS ด้วยการพิจารณาความถี่ในการเช็คอินของผู้ใช้ร่วมกับค่าความชอบของผู้ใช้จากการคำนึงถึงการกลับมาเยี่ยมชมหรือใช้บริการซ้ำต่อสถานที่ของผู้ใช้ 2. การจัดอันดับสถานที่ที่น่าสนใจ จากค่าคะแนนที่มากที่สุดของเซทรายการสถานที่ด้วยการพิจารณาความถี่ในการเช็คอินของผู้ใช้และเซทรายการสถานที่ด้วยการพิจารณาค่าความชอบของผู้ใช้โดยใช้การคำนวณแบบวนซ้ำ และ 3. การสร้างรายการแนะนำสถานที่ที่น่าสนใจ โดยระบบจะทำการเลือกสถานที่ที่น่าสนใจมากที่สุด N อันดับที่ตั้งอยู่ระแวกใกล้เคียงกับที่อยู่ปัจจุบันของผู้ใช้จากรายการแนะนำสถานที่ที่ถูกจัดอันดับไว้จากขั้นตอนก่อนหน้า จากผลการทดลองที่นำเสนอโดยใช้ฐานข้อมูลการเช็คอินของผู้ใช้จาก Foursquare ใน 5 เมือง คือ Chiyoda Minato Shinjuku Shibuya และ Chuo สำหรับการจำลองที่อยู่ปัจจุบันของผู้ใช้เพื่อสร้างรายการแนะนำสถานที่ที่น่าสนใจเฉพาะพื้นที่ได้ใช้การกำหนดของเขตของผู้ใช้ที่เรียกว่า Minimum Bounding Rectangles (MBRs) และทำการเปรียบเทียบกับขั้นตอนวิธี HITS โดยขั้นตอนวิธีการที่นำเสนอ NILR สามารถมีความถูกต้องในการสร้างรายการแนะนำสถานที่ที่น่าสนใจโดยพิจารณาจากค่าเฉลี่ยของรายการแนะนำขนาด 1 สถานที่จนถึง 10 สถานที่ที่ดีกว่าขั้นตอนวิธี HITS based model ที่ 4.49% และ 3.05% โดยใช้วิธี Precision และ Recall ตามลำดับและ NILR มีความถูกต้องในการจัดอันดับที่ดีกว่าขั้นตอนวิธี HITS ที่ 33.09% โดยใช้วิธี Average ranking ตามลำดับ

บรรณานุกรม

Zheng, Y., Zhang, L., Xie, X., & Ma, Y. (2009) Mining interesting locations and travel sequences from gps trajectories. In WWW, pages 791--800, 2009.

Xiao, X., Zheng, Yu., Luo, Q., & Xie, X. (2010) Finding similar users using category-based location history. In Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems (GIS '10). Association for Computing Machinery, New York, NY, USA, 442–445.

Zheng, W. V., Cao, B., Zheng, Y., Xie, X. & Yang, Q. (2010) Collaborative filtering meets mobile recommendation: a user-centered approach. In Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence (AAAI'10). AAAI Press, 236–241.

Zheng, W. V., Zheng, Y., Xie, X., & Yang, Q. (2010) Collaborative location and activity recommendations with GPS history data. In Proceedings of the 19th international conference on World wide web (WWW '10).

Zheng, Y., & Xie, X. (2011) Location-based social networks: Locations. In Computing with Spatial Trajectories, Springer, 277--308.

Symeonidis, P., Papadimitriou, A., Manolopoulos, Y., Senkul, P., & Toroslu, I. (2011) Geo-social recommendations based on incremental tensor reduction and local path traversal. In Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Location-Based Social Networks (LBSN '11). Association for Computing Machinery, New York, NY, USA, 89–96.

Zhou, D., Wang, B., Rahimi, S. M., & Wang, X. (2012). A Study of Recommending Locations on Location-Based Social Network by Collaborative Filtering. Lecture Notes in Computer Science, 255–266.

Levandoski, J. J., Sarwat, M., Eldawy, A., & Mokbel, F. M. (2012). LARS: A Location-Aware Recommender System. 2012 IEEE 28th International Conference on Data Engineering, Washington, DC, 2012, pp. 450-461.

Bao, J., Zheng, Y., & Mokbel, F. M. (2012) Location-based and preference-aware recommendation using sparse geo-social networking data. In Proceedings of the 20th

International Conference on Advances in Geographic Information Systems (SIGSPATIAL '12). ACM, New York, NY, USA, 199-208.

Long, X. & Joshi, J. (2013) A HITS-based POI recommendation algorithm for location-based social networks. In Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM '13). ACM, New York, NY, USA, 642-647.

Yuan, Q., Cong, Gao., Ma, Z., Sun, A., & Magnenat- Thalmann, N. (2013) Time-aware point-of-interest recommendation. In Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval (SIGIR '13), 2013.

Gao, H., Tang, J., Hu, X., & Liu, H. (2013) Exploring temporal effects for location recommendation on location-based social networks. In Proceedings of the 7th ACM conference on Recommender systems (RecSys '13), 2013.

Preotiuc-Pietro, D. & Cohn, T. (2013) Mining user behaviours: a study of check-in patterns in location based social networks. In Proceedings of the 5th Annual ACM Web Science Conference (WebSci '13), 2013.

Yin, H., Sun, Y., Cui, B., Hu, Z. & Chen, L. (2013) LCARS: a location-content-aware recommender system, Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining, August 11-14, 2013.

Liu, X., Liu, Y., Aberer, K., & Miao, C. (2013) Personalized point-of-interest recommendation by mining users' preference transition. In Proceedings of the 22nd ACM international conference on Information & Knowledge Management (CIKM '13), 2013.

Zheng, Y., Capra, L., Wolfson O., & Yang, H. (2014) Urban computing: Concepts, methodologies, and applications. ACM Transactions on Intelligent Systems and Technology 5, 3 (2014), 38--55.

Wang, H., Li, G., & Feng, J. (2014) Group-Based Personalized Location Recommendation on Social Networks. In: Chen L., Jia Y., Sellis T., Liu G. (eds) Web Technologies and Applications. APWeb 2014. Lecture Notes in Computer Science, vol 8709. Springer.

Yu, Y. & Chen, X. (2015) A survey of point-of-interest recommendation in location-based social networks. In Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence, 2015.

Zhang J-D., Chow, C-Y., & Zheng, Y. (2015) ORec: An Opinion-Based Point-of-Interest Recommendation Framework. In Proceedings of the 24th ACM International on Conference on Information and Knowledge Management (CIKM '15), 2015.

Yang, D., Zhang, D., Zheng, W. V., & Yu, Z. (2015) Modeling User Activity Preference by Leveraging User Spatial Temporal Characteristics in LBSNs, in IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 45, no. 1, pp. 129-142.

Bagci, H. & Karagoz, P. (2016) Context-aware location recommendation by using a random walk-based approach. Knowl. Inf. Syst. 47, 2 (May 2016), 241-260.

Zhao, S., King, I, & Lyu, R. M. (2016) A survey of point-of-interest recommendation in location-based social networks. arXiv preprint arXiv:1607.00647, 2016.

Chen, J., Li, X., Cheung, K. W., & Li, K. (2 0 1 6) Effective successive POI recommendation inferred with individual behavior and group preference. Neurocomput. 210, C (October 2016), 174–184.

Zhang., J., & Chow C. (2016) TICRec: A Probabilistic Framework to Utilize Temporal Influence Correlations for Time-Aware Location Recommendations, in IEEE Transactions on Services Computing, vol. 9, no. 4, pp. 633-646, 1 July-Aug, 2016.

Ying, Y., Chen, L., & Chen, G. (2017) A temporal-aware poi recommendation system using context-aware tensor decomposition and weighted hits. Neurocomputing, 242 (2017), pp. 195-205.

Si, Y., Zhang, F., Liu, W. (2017) CTF-ARA: An adaptive method for POI recommendation based on check-in and temporal features. Knowl.-Based Syst. 128, 59–70, 2017.

Kefalas, P., Symeonidis, P., & Manolopoulos, Y. (2017) Recommendations based on a heterogeneous spatio-temporal social network, World Wide Web (2017) 1– 27, 2017.

Ding, Z., Li, X., Jiang, C., & Zhou, M. (2018). Objectives and state-of-the-art of location-based social network recommender systems, ACM Comput. Surv. (CSUR), 51 (1) (2018).

Lim, H. K., Chan, J., Leckie, C., & Karunasekera, J. (2018) Personalized trip recommendation for tourists based on user interests, points of interest visit durations and visit recency. Knowl Inf Syst 54, 375–406 (2018).

Ding, Z., Li, X., Jiang, C., & Zhou, M. (2018) Objectives and State-of-the-Art of Location-Based Social Network Recommender Systems. ACM Comput. Surv. 51, 1, Article 18, 2018.

Cao, J., Xu, S., Zhu, X., Lv, R. & Liu, B. (2018) Effective fine-grained location prediction based on user check-in pattern in LBSNs, Journal of Network and Computer Applications, Volume 108, 2018, Pages 64-75.

Puspitaningrum, D., Fernando, J., Afriando, E., Utama, P. F., Rahmadini, Rina & Pinata, Y. (2019) Finding Local Experts for Dynamic Recommendations Using Lazy Random Walk, 7th International Conference on Cyber and IT Service Management (CITSM), Jakarta, Indonesia, 2019, pp. 1-6

Lim, H. K., Chan, J., Karunasekera, J., & Leckie, C. (2019) Tour recommendation and trip planning using location-based social media: a survey. Knowl Inf Syst 60, 1247–1275 (2019).

Geng, B., Jiao, L., Gong, M., Li, L. & Wu, Y. (2019) A two-step personalized location recommendation based on multi-objective immune algorithm, Information Sciences, Volume 475, 2019, Pages 161-181.

Xu, S., Fu, X., Cao, J., Liu, B. & Wang, Z. (2020) Survey on user location prediction based on geo-social networking data. World Wide Web 23, 1621–1664.